Is Peer Review a Good Idea?*

Remco Heesen†‡ Liam Kofi Bright§

September 19, 2018

Abstract

Pre-publication peer review should be abolished. We consider the effects that such a change will have on the social structure of science, paying particular attention to the changed incentive structure and the likely effects on the behavior of individual scientists. We evaluate these changes from the perspective of epistemic consequentialism. We find that where the effects of abolishing pre-publication peer review can be evaluated with a reasonable level of confidence based on presently available evidence, they are either positive or neutral. We conclude that on present evidence abolishing peer review weakly dominates the status quo.

---

*Both authors contributed equally. Thanks to Justin Bruner, Adrian Currie, Cailin O'Connor, and Jan-Willem Romeijn for valuable comments. RH was supported by an Early Career Fellowship from the Leverhulme Trust and the Isaac Newton Trust. LKB was supported by NSF grant SES 1254291.

†Department of Philosophy, School of Humanities, University of Western Australia, Crawley, WA 6009, Australia. Email: remco.heesen@uwa.edu.au.

‡Faculty of Philosophy, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, UK.

§Department of Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London WC2A 2AE, UK. Email: liamkbright@gmail.com.

1 Introduction

Peer review plays a central role in contemporary academic life. It sits at the critical juncture where scientific work is accepted for publication or rejected. This is particularly clear when the results of scientific work are communicated to non-scientists, e.g., by journalists. The question "Has this been peer reviewed?" is commonly asked, and a positive answer is frequently taken to be a necessary and sufficient condition for the results to be considered serious science.

Given these circumstances, one might expect peer review to be an important topic in the philosophy of science as well. Peer review should arguably play a more prominent role in the debate about demarcation criteria (what separates science from other human pursuits?), as it seems to be used in practice exactly to differentiate scientific knowledge from other claims to knowledge, at least by journalists. Yet as far as we know, social-procedural accounts of science, like the one found in Longino (1990), remain in the minority and usually do not place great emphasis on peer review in particular. Aside from this particular debate, there are normative questions about the proper epistemic role of peer review and more practical questions about the extent to which it manages to fulfill them, all of which should interest philosophers of science.

But there has been surprisingly little work on peer review by philosophers of science. Most of what exists has focused on the role of biases in peer review, see for example Saul (2013, §2.1), Lee et al. (2013), Jukola (2017), Katzav and Vaesen (2017), and Heesen (2018). We are not aware of any philosophical discussion of the strengths and weaknesses of peer review as such (the above examples presuppose its overall legitimacy by discussing its implementation). Some work along these lines does exist outside of philosophy, either in the form of opinion pieces (Gowers 2017) or occasionally full-length articles (Smith 2006). Such work tends to be vague about the normative standard against which peer review or its alternatives are to be evaluated, something we aim to remedy in section 2.

Here we bring together the work of philosophers of science (especially social epistemologists of science) who have written about the strengths and weaknesses of various aspects of the social structure of science and empirical work about the effects of peer review. We argue that where philosophers of science have claimed the social structure of science works well, their arguments tend to rely on things other than peer review, and that where specific benefits have been claimed for peer review, empirical research has so far failed to bear these out. Comparing this to the downsides of peer review, most prominently the massive amount of time and resources tied up in it, we conclude that we might be better off abolishing peer review.

Some brief clarifications. Our target is pre-publication peer review, that is the review of a manuscript intended for publication, where publication is withheld until one or more editors deem the manuscript to have successfully passed peer review. We set aside other uses of peer review (e.g., of grant proposals or conference abstracts) and we explicitly leave room for post-publication peer review, where manuscripts are published before review. Because of this last point, some readers may think that our terminology ('abolishing pre-publication peer review') suggests a more dramatic change than what we actually advocate. We invite such readers to substitute in their preferred terminology. We should also clarify that we use 'science' in a broad sense to include the natural sciences, the social sciences, and the humanities.

The overall structure of our argument is as follows. We think there are a number of clear benefits to abolishing pre-publication peer review. In contrast, while various benefits of the existing system (downsides of abolishing peer review) have been suggested, we do not think there exist any that have clear empirical support. Insofar as empirical research exists, it is ambiguous in some cases, and speaks relatively clearly against the claimed benefit of the existing system in others. While we admit to a number of cases where the evidence is ambiguous or simply lacking (see especially section 5), we claim that the present state of the evidence suggests that abolishing pre-publication peer review would lead to a Pareto improvement: each factor considered is either neutral or favors our proposal.

Our primary aim here is to evaluate the current system, but we believe that is only really possible by comparing it to an alternative. We are not claiming that the proposal we put forward is the best of all possible alternatives. It has been constructed to be a system which could constitute a Pareto improvement over the current system. Given that it has not actually been implemented yet, we cannot guarantee it would work as advertised or what empirical properties it would have. But in offering a relatively specific alternative, we hope to get people thinking about real change, which pointing out problems with the present system has so far failed to do.

Even despite this proviso, we realize that ours is a strong claim, and our proposal a large change to the social structure of science. It is therefore important to highlight that our central claim concerns the balance of presently available evidence. We are not further claiming that the matter is so conclusively settled as to render further research superfluous or wasteful. On the contrary, we think there are a number of points in our argument where the presently available evidence is severely limited, and we take the calls for further empirical research we make in those places to be just as important a part of the upshot of our paper as our positive proposal. We hope, therefore, that even a skeptical reader will read on; if not to be convinced of the need of abolishing pre-publication peer review, then at least to see where in our view their future research efforts should concentrate if they are to shore up pre-publication peer review's claims to good epistemic standing.

2 Setting the Stage

The purpose of peer review is usually construed in terms of quality control. For example, Katzav and Vaesen (2017, 6) write "The epistemic role of peer review is assessing the quality of research", and this seems to be a common sentiment per, e.g., Eisenhart (2002, 241) and Jukola (2017, 125). But how well does peer review succeed in its purpose of quality control? The empirical evidence (reviewed below) is mixed at best. As one prominent critic puts it, "we have little evidence on the effectiveness of peer review, but we have considerable evidence on its defects" (Smith 2006, 179).

Peer review's limited effectiveness would perhaps not be a problem if it required little time and effort from scientists. But in fact the opposite is true. Going from a manuscript to a published paper involves many hours of reviewing work by the assigned peer reviewers and a significant time investment from the editor handling the submission. The editor and reviewers are all scientists themselves, so the epistemic opportunity cost of their reviewing work is significant: instead of reviewing, they could be doing more science.

Given these two facts—high (epistemic) costs and unclear benefits—we raise the question whether it might be better to abolish pre-publication peer review. In the following we provide our own survey and assessment of the evidence that bears on this question. Our conclusions are not sympathetic to peer review. However, we encourage any proponents of peer review to give their own assessment. We only ask that any benefits claimed for peer review are backed up by empirical research, and that they are epistemic benefits, i.e., we ask for empirical evidence that peer review makes for better science on science's own terms.

We take the status quo to be as follows. The vast majority of scientific work is shared through journal publications, and the vast majority of journals uses some form of pre-publication peer review. Ordinarily this means that an editor assigns one to three peers (scientists whose expertise intersects the topic of the submission), who provide a report and/or verdict on the submission's suitability for publication. The peer reviews feed into the final judgment: the submission is accepted or rejected with or without revisions.

Our proposal is to abolish pre-publication peer review. Scientists themselves will decide when their work is ready for sharing. When this happens, they publish their work online on something that looks like a preprint archive (although the term "preprint" would not be appropriate under our proposal). Authors can subsequently publish updated versions that reply to questions and comments from other scientists, which may have been provided publicly or privately. Most journals will probably cease to exist, but the business of those that continue will be to create curated collections of previously published articles. Their process for creating these collections will presumably still involve peer review, but now of the post-publication variety.

Our proposal is in line with how certain parts of mathematics and physics already work: uploading a paper to arXiv is considered publishing it for most purposes, with journal peer review and publication happening almost as an afterthought (Gowers 2017). Indeed, journal publication can function as something like a prize, accruing glory to the scientist who achieves it but doing little to actually help spread or diffuse the idea beyond calling attention to something that has already been made public elsewhere. We are not aware of any detailed comparative studies of the effects these changes have had in those fields, so we will not rest any significant part of our argument on this case. But for those who worry that science will immediately and irrevocably fall apart without peer review, we point out that this does not appear to have happened in the relevant parts of mathematics and physics.

In the remainder of this paper we break down the consequences of our proposal. Our strategy here is to focus on a large number (hopefully all) aspects of the social structure of science that will be affected. In particular, the reader may already have a particular objection against our proposal in mind. We encourage such a reader to skip ahead to the section where this objection is discussed before reading the rest of the paper.

For example, one reader may think that peer review as currently practiced is important because it forces scientists to read and review each other's work, and without peer review they will spend less time on such tasks. This is discussed in section 3.2. Another reader may worry that without peer review and the journal publications that go with them it will be more difficult to evaluate scientists for hiring or promotion (section 3.5). Yet another reader may be concerned about losing peer review's ability to prevent work of little merit from being published, or at least to sort papers into journals by epistemic merit so scientists can easily find good work (section 4.1). A fourth reader might think peer review plays an important role in detecting fraud or other scientific malpractice (section 4.2). A fifth reader may think the guarantee provided to outsiders when something has been peer reviewed is an important reason to preserve the status quo (section 5.1). And a sixth reader may want to point out that anonymized peer review gives relatively unknown scientists a chance at an audience by publishing in a prestigious journal, whereas on our proposal perhaps only antecedently prominent scientists will have their work read and engaged with (section 5.2).

Other aspects of the social structure of science that will be considered: whether and when scientists share their work (section 3.1), how many papers are published by women or men (section 3.3), library resources (section 3.4), the power of editors as gatekeepers (section 3.6), science's susceptibility to fads and fashions (section 4.3), and ways to get credit for scientific work other than through journal publications (section 4.4). In each case we evaluate whether the net effects of our proposal on that aspect can be expected to be positive. To tip our hand: aspects where we will claim a benefit are gathered in section 3, aspects where we expect little or no change are in section 4, and aspects that we consider neutral due to a present lack of evidence are in section 5.

In making these evaluations, we commit to a kind of epistemic consequentialism (cf. Goldman 1999). One may think of what we are doing as roughly analogous to the utilitarian principle, where for each issue our yardstick is whether pre-publication peer review shall generate the greatest amount of knowledge produced in the least amount of time. More specifically, we consider changes in the incentive structure and expected behaviour of scientists, as well as other changes that would result from abolishing pre-publication peer review. We evaluate these changes in terms of their expected effect on the ability of the scientific community to produce scientific knowledge in an efficient manner. Working out in detail what such an epistemic consequentialism would look like would be very complicated, and we do not attempt the task here. For most of the issues we consider, we think that the calculus is sufficiently clear that fine details do not matter. Where it is unclear (the issues discussed in section 5) we think this results from ignorance of empirical facts about the likely effect of policies, rather than conceptual unclarity in the evaluative metric. So we do not need to use our consequentialist yardstick to settle any difficult tradeoffs. All we need for our purposes is to make it clear that we are evaluating the peer review system by how well it does in incentivizing efficient knowledge production.

What do we mean by the incentive structure of science, mentioned in the previous paragraph? This addresses the motivations of scientists. Scientists are rewarded for their contributions with credit, i.e., with recognition from their peers as expressed through such things as awards, citations, and prestigious publications (Merton 1957, Hull 1988, Zollman 2018). Scientific careers are largely built on the reputations scientists acquire in this way (Latour and Woolgar 1986, chapter 5). As a result, scientists engage in behaviors that improve their chances of credit (Merton 1969, Dasgupta and David 1994, Zollman 2018).

While individual scientists may be motivated by credit to different degrees (curiosity, the thrill of discovery, and philanthropic goals are important motivations for many as well), the effect on careers means that credit-maximizing behavior is to some extent selected for. Thus we think it important to ensure that our proposal does not negatively affect the incentives currently in place for scientists to work effectively and efficiently.

3 Benefits of Abolishing Peer Review

3.1 Sharing Scientific Results

An important feature of (academic) science is that there is a norm of sharing one's findings with the scientific community. This has been referred to as the communist norm (Merton 1942). In recent surveys, scientists by and large confirm both the normative force of the communist norm and their actual compliance (Louis et al. 2002, Macfarlane and Cheng 2008, Anderson et al. 2010). This norm is epistemically beneficial to the scientific community, as it prevents scientists from needlessly duplicating each other's work.

Will abolishing peer review affect this practice? In order to answer this question, we need to know what motivates scientists to comply with the communist norm, that is to share their work. On the one hand there is the feeling that they ought to share generated by the existence of the norm itself. There is no reason to expect this to be changed by abolishing peer review.

On the other hand there is the motivation generated by the desire for credit. According to the priority rule, the first scientist to publish a particular discovery gets the credit for it (Merton 1957, Dasgupta and David 1994, Strevens 2003). So a scientist who wants to get credit for her discoveries has an incentive to publish them as quickly as possible, in order to maximize her chances of being first. Recent work suggests that this applies even in the case of smaller, intermediate discoveries (Boyer 2014, Strevens 2017, Heesen 2017b). All of this helps motivate scientists to share their work.

If peer review were to be abolished, the communist norm and the priority rule would still be in effect, so scientists would still be motivated to share their work as quickly as possible. However, the following change would occur.

In the absence of pre-publication peer review, scientists would be able to share their discoveries more quickly. In the current system, peer review can hold up publication for significant amounts of time, especially in the case of fields with high rejection rates or long turnaround times. During this time, other scientists cannot build on the work and may spend their time needlessly duplicating the work. Cutting out this lag by letting scientists publish their own work when they think it is ready will speed up scientific progress. While being faster is not always better (it may increase the risk of error, cf. Heesen 2017c), in this case delays in publication are reduced without any reduction in the time spent on the scientific work itself.

To some extent this already occurs. Scientists, especially well-connected scientists, already share preprints that make the community aware of their work in advance of publication. For people who regularly do this, practically speaking little would change upon adopting the system we advocate. However, our proposal turns pre-journal-publication dispersal of work from a privilege of a well-connected few into the norm for everyone.

On this point, then, abolishing peer review is a net positive, as scientists will still be incentivized to share their work as soon as possible, but the delays associated with pre-publication peer review are removed.

3.2 Time Allocation

The current system restricts the way scientists are allowed to spend their time. For each paper submitted to a journal, a number of scientists are conscripted into reviewing it, and at least one editor has to spend time on that paper as well.

On our proposal, scientists would be free to choose how much of their time to spend reading and reviewing others' work as compared to other scientific activities. Some scientists would spend less time reviewing, some scientists would spend more, and some would spend exactly as much as under the current system.

For scientists in the latter category our proposal makes no difference, while for scientists in the other two categories our proposal represents a net improvement of how they spend their time, at least in their own judgments. We think people are the best judges of how to use their own time and labor.

We thus trust scientists' decisions in these regards, and welcome changes that would render many scientists' choices about how to allocate their own labor independent of the preferences of the relatively small number of editorial gatekeepers.

So we assume that scientists are well-placed to judge how best to use their own abilities to meet the community's epistemic needs. We claim, moreover, that the reward structure of science is set up so as to make it in their interest to do so: the credit economy incentivizes scientists to spend their time on pursuits the epistemic value of which will be recognized by the community (Zollman 2018). Hence freeing up the way scientists allocate their time leads to net epistemic benefits to the scientific community.

One might object that journals perform a useful epistemic sorting role, telling scientists what is worth spending their time on. We will address these concerns in section 4.1.

One might think that this would lead scientists to spend significantly less time reading and reviewing others' work. If this is right, we still think it would be an overall improvement for the reasons mentioned above. But we also want to point out that this is not as obvious a consequence as it may seem. Here are two reasons to expect scientists to spend as much time or more reading and reviewing on our proposal. First, for many scientists reading and reviewing are intrinsically valuable and can help their own research. Second, the current system provides no particular incentive to read and review either: scientists agree to review only because they independently want to or because they feel an obligation to the research community. While no one scientist is conscripted, at the group level editors are going to keep going until they find someone. This can amount to picking whomever is most weak-willed or under some extra-epistemic social pressure. It is not obvious that this way of deciding who does the reviewing has much to recommend it. Any rewards that exist for reviewing will still exist on our proposal, and may be amplified by the possibility of making post-publication reviews public.

3.3 Gender Skew in Publications

Male scientists publish more, on average, than female scientists, a phenomenon known as the productivity puzzle or productivity gap (Zuckerman and Cole 1975, Valian 1999, Prpić 2002, Etzkowitz et al. 2008). Several explanations have been suggested, none of which are entirely satisfactory (see especially Etzkowitz et al. 2008, 409–412). Two of these explanations that are relevant to our concerns here are the direct effects of gender bias and the indirect effects of the expectation of gender bias.

There is some evidence of gender bias in peer review, although this is not unambiguous (see Lee et al. 2013, 7–8, Lee 2016, and references therein). Insofar as there is gender bias—in the sense of women's work being judged more negatively by peer reviewers—abolishing peer review will remove this and help level the playing field for men and women. We expect positive epistemic consequences from the removal of these arbitrarily different standards.

While the evidence of gender bias in peer review is not entirely clear-cut, there is good evidence that women *expect* to face gender bias in peer review (see Lee 2016, Bright 2017b, Hengel 2018, and references therein). In an effort to overcome this perceived bias, women tend to hold their own work to higher standards. Hengel (2018) provides evidence that women spend more time correcting stylistic aspects of their paper during peer review, presumably due to higher expectations of scrutiny on such apparently superficial elements of their work. On the plausible assumption that if women have higher standards for each paper they will produce fewer papers overall, this means that the mere expectation of gender bias can contribute to the productivity gap.

After abolishing peer review both women and men will hold their work primarily to their own individual standards of quality, and secondarily to their expectations of the response of the entire scientific community, but not to their expectations of the opinion of a small arbitrary group of gatekeepers. We do not know whether this will lead the women to behave more like the men (producing more papers) or the men to behave more like the women (holding individual papers to a higher standard of quality). However, in line with our view above that scientists are well-placed to judge how best to spend their own time, we take it that any resulting change in behavior will be a net epistemic positive.

3.4 Library Resources

Journal subscription fees currently take up a large amount of library resources (RIN 2008, Van Noorden 2013). To summarize some key figures from the 2008 report: research libraries in the UK spent between £208,000 and £1,386,000 on journal subscriptions annually (and that was a decade ago, with subscriptions having risen substantially since). The cost for publishing and distributing a paper was estimated to be about £4,000, or about £6.4 billion per year in total. Savings from moving to author-paid open access were estimated at £561 million, about half of which would directly benefit libraries.

On our proposal, this is replaced by the cost of maintaining one or more online archives of scientific publications. The example of existing large preprint archives like arXiv and bioRxiv suggests that maintaining such archives can be done at a fraction of the cost currently spent on journal subscription fees. As a rough guideline, Van Noorden (2013) estimates maintenance costs of arXiv at just $10 per article. So our proposal involves significant savings on library resources, which could be used to expand collections, retain more or better trained staff, or other purposes that would be of epistemic benefit to the scientific community.

Two additional effects should be considered in relation to this. First, the fact that the online archive will be open access means that scientific publications will be available to everyone, not just to those with a library subscription or some other form of access to for-profit scientific journals.

Second, the fact that any value added by for-profit journals would be taken away. The two tasks currently carried out by journals that could plausibly be supposed to add value to scientific publications are peer review and copy-editing (Van Noorden 2013). It is the purpose of all other sections of this paper to argue that peer review does not in fact (provably) add value, so we set that aside. This leaves copy-editing. We propose that libraries use some of the funds freed up from journal subscriptions to employ some copy-editors. Each university library would make copy-editing services available to the scientists employed at that university. We contend that, after paying for the maintenance of an online archive and a team of copy-editors, under our proposal libraries would still end up with more resources for other pursuits than under the current system.

We note that this particular advantage of our proposal is a bit more historically contingent than the others. There seems to be no particular reason why pre-publication peer review has to be implemented through for-profit journals, and if the open access movement has its way we might be able to free up these library resources without abolishing pre-publication peer review. But our proposal also achieves this goal, and so we count it as an advantage relative to the system as it is currently actually implemented.

3.5 Scientific Careers

The 'publish or perish' culture in science has been widely noted (e.g., Fanelli 2010). Universities judge the research productivity of scientists through their publications in (peer reviewed) journals, with some focusing more on 'quantity' (counting publications) and others on 'quality' (publishing in prestigious journals). Scientific journals and the system of pre-publication peer review thus play an important role in shaping scientific careers. What will become of this if peer review is abolished?

We note first that the 'publish or perish' culture is a subset of a larger system which we discussed above: the credit economy. Publishing in a journal is one way to receive credit for one's work, but there are others, most prominently citations and awards. Scientific careers depend on all of these, with different institutions weighting quantity of publications, quality of publications, citation metrics, and awards and other honors differently.

Any of these types of credit represents some kind of recognition of the scholarly contributions of the scientist by her peers. But here we distinguish two types of credit, which we will call short-run credit and long-run credit. Getting a paper through peer review yields a certain amount of credit: more for more prestigious journals, less for less prestigious ones. But this is short-run credit in the following sense. The editor and the peer reviewers judge the technical adequacy and the potential impact of the paper, shortly after it is written. Their judgment is essentially a prediction of how much uptake the paper is likely to receive in the scientific community.

In contrast, citations (as well as awards, prizes, inclusion in anthologies or textbooks, etc.) represent long-run credit. They *are* the uptake the paper receives in the scientific community. Long-run credit is both a more considered opinion of the scientific importance of the paper and a more democratic one (citations can be made by anyone, and awards usually reflect a consensus in the scientific community, whereas peer review is normally done by up to three individuals). So long-run credit reflects a more direct and better estimate of the real epistemic value of a contribution to science.

So what would the effect of our proposal be? For better or worse, our proposal does not make it impossible for universities to use metrics to judge research productivity. While journal rankings and impact factors would disappear, citation metrics for individual scientists and papers would still be available. This may mean that universities stop judging their scientists based on the impact factors of the journals they publish in and start judging them on the actual citation impact of their papers. More generally, our proposal will decrease or remove the role of short-run credit in shaping career outcomes and increase the role of long-run credit, which we take to be a better measure of scientific importance. So we think this is an improvement on the status quo.

What about junior hires and related career decisions, where long-run credit may be absent or minimal? If abolishing peer review means completely getting rid of journals and the associated prestige rankings, this robs hiring departments of some information regarding the scientific importance of candidates' work. If this means those on the hiring side need to read and form an opinion of candidates' work for themselves, we do not think that is a bad thing. This would of course take time, but if journals and peer review are completely abolished, that just means the time spent reviewing the paper is transferred to the people considering hiring the scientist, which again, we do not think is a bad thing. In fact, since very few academics are on a hiring committee year after year, whereas referee requests are a constant feature while one is in the community, we think that even this added burden when hiring might still be a net time-saver for academics.

But it does not have to be that way. We never said journals and peer review have to be completely abolished—our proposal in section 2 explicitly suggests journal issues may still appear, but as curated collections of articles based on post-publication peer review. So short-run credit based on journal prestige need not disappear. It need not even be slower as there is no particular reason post-publication peer review needs to take longer than pre-publication peer review. But there is the added advantage that the paper is already published while it undergoes peer review, so the wider community outside the assigned reviewers also has a chance to respond before it is included in a journal.

3.6 The Power of Gatekeepers

The discussion immediately above touched on another effect, one that we think is worth bringing out as a benefit of our proposal in its own right. As mentioned our proposal suggests that in evaluating the importance of scientific work we decrease our reliance on short-run credit (journal prestige), with a corresponding increase in long-run credit (citations, among other things).

This means that the overall credit associated with a particular paper depends less on the judgments made by an editor and a small number of reviewers, and more on its actual uptake in the larger scientific community.

Editors in particular currently play a large role in determining which scientific work is worthy of attention, as they are a relatively small group of people with a deciding vote in the peer review process of a large number of papers. They are often referred to as gatekeepers for this reason (Crane 1967). Our proposal entails significantly decreasing both the prevalence and importance of this role. By replacing some of this importance with long-run credit, which comes from the scientific community as a whole, it makes the evaluation of scientific work a more democratic process. Not only is there some reason to think that democratic evaluation of scientific claims is more in line with general communal norms accepted within science (Bright et al. 2018), but general arguments from democratic theory and social epistemology of science give epistemic reason to welcome the increased independence of judgment and evaluation this would introduce (List and Goodin 2001, Heesen et al. forthcoming, Perović et al. 2016, 103–104).

4 Where Peer Review Makes No Difference

In this section we consider a number of aspects of the scientific incentive structure for which we think a case can be made that abolishing peer review will leave them basically unaffected. This serves partially to forestall objections to our proposal that we anticipate from defenders of the peer review system, and partially to avoid overstating our case—in some of what follows we argue that abolishing peer review will likely have no effect in cases where one might have expected it to be beneficial.

4.1 Epistemic Sorting

Given the stated purpose of peer review mentioned in section 2 the first and most apparent disadvantage of our proposal is that it would remove the epistemic filter on what enters into the scientific literature. One might worry that the scientific community would lose the ability to maintain its own epistemic standards, and thus the general quality of scientific research would be reduced. We argue here that despite the intuitive support this idea might have, the present state of the literature on scientific peer review does not support it.

Separate out two kind of epistemic standards one may hope that the peer review system maintains. First, that peer review allows us to identify especially meritorious work and place it in high profile journals, while ensuring that especially shoddy work is kept from being published. Call this the 'epistemic sorting' function of peer review. Second, that peer review allows for the early detection of fraudulent work or work that otherwise involves research misconduct. Call this the 'malpractice detection' function of peer review. We deal with each of these in turn.

Let us step back and ask why, from the point of view of epistemic consequentialism, one would want peer review to do any sort of epistemic sorting. We take the answer to be that epistemic sorting helps scientists fruitfully direct their time and energy by selecting the best work and bringing it to scientists' attention through publication in journals. They read and respond to that which is most likely to help them advance knowledge in their field.

How could peer review achieve this? One might hope that peer review functions by keeping bad manuscripts out of the published literature and letting good manuscripts in. This, however, is a non-starter. There are far too many journals publishing far too many things, with standards of publication varying far too wildly between them, for the sheer fact of having passed peer review somewhere to be all that informative as to the quality of a manuscript.

Instead, if peer review is to serve anything like this purpose it must be because reviewers are able (even if imperfectly) to discern the relative degree of scientific merit of a work, and sort it into an appropriately prestigious journal. Epistemic sorting happens not via the binary act of granting or withholding publication, but rather through sorting manuscripts into journals located on a prestige hierarchy that tracks scientific merit.

A necessary condition for epistemic sorting to work as advertised is that reviewers be reliable guides to the merit of the scientific work they review. Our first critique is that this necessary condition does not seem to be met. Investigation into reviewing practices has not generally found much inter-reviewer reliability in their evaluations (Peters and Ceci 1982, Ernst et al. 1993, Lee et al. 2013, 5–6). What this means is that one generally cannot predict what one reviewer will think of a manuscript by seeing what another reviewer thought. If there was some underlying epistemic merit scientists were accurately (even if falteringly) discerning by means of their reviews, one would expect there to be correlations in reviewers evaluations. However, this is not what we find. Indeed, one study of a top medical journal even found that "reviewers... agreed on the disposition of manuscripts at a rate barely exceeding what would be expected by chance" (Kravitz et al. 2010, 3). Findings like these are typical in the literature that looks at inter-reviewer reliability (for a review of the literature see Bornmann 2011, 207). The available evidence does not provide much support for the idea that pre-publication peer review detects the presence of some underlying quality.

Our second critique of the epistemic sorting idea speaks more directly to the ideal it tracks. We are not persuaded that the best way to direct scientists' attention is to continually alert them to the best pieces of individual work, and have them proportion their attention according to position on a prestige hierarchy. We take it the intuition behind this is a broadly meritocratic one. This intuition has been challenged by some modeling work (Zollman 2009). While Zollman retained some role for peer review, his model still found that striving to select the best work for publication is not necessarily best from the perspective of an epistemic community; his model favored a greater degree of randomization.

We do not wish to rest our case on the results of one model which in any case does not fully align with our argument, but it highlights that the ideal of meritocracy stands in need of more defense than it is typically given. We take it that scientists most fruitfully direct their attention to that package of previous work and results which, when combined, provides them with the sort of information and perspectives they need to best advance their own epistemically valuable projects. It is a presently undefended assumption that this package of work should be composed of works which are themselves individually the most meritorious work, or that paying attention to the prestige hierarchy of journals and proportioning one's attention accordingly will be useful in constructing such a package. Hence, even if it did turn out that the peer review system could sort according to scientific merit, it is an underappreciated but important fact that this is not the end of the argument. Further defense of the purpose of this kind of epistemic sorting is needed from the point of view of epistemic consequentialism.

Before moving on we note a potential objection. Even if one did not think that peer review was detecting some underlying quality or interestingness, one might think that the process of feedback and revision which forms part of the peer review system would be beneficial to the epistemic quality of the scientific literature. In this way epistemic sorting may have a positive epistemic effect even if it fails in its primary task.

However, this returns us to the points regarding gatekeepers and time allocation from section 3. We are not opposed to scientists reading each other's work, offering feedback, and updating their work in light of that. This can indeed lead to improvements (Bornmann 2011, 203), though in this context it is worth noting the results of an experiment in the biomedical sciences, which found that attempting to attach the allure of greater prestige to more epistemically high caliber work did little to actually improve the quality of published literature (Lee 2013). Fully interpreting these results would require discussion of the measures of quality used in such literature. We do not intend to do that here, since we do not intend to dispute the point that it is desirable for scientists to give feedback and respond to it.

We would expect this sort of peer-to-peer feedback to continue under a system without pre-publication peer review. Curiosity, informal networking, collegial responsibilities, and the credit incentives to engage with others' work and make use of new knowledge before others do; these would all be retained even without pre-publication peer review. What would be eliminated is the assignment of reviewing duties to papers that scientists did not independently decide were worth their time and attention, and the necessity of giving uptake to criticism (in order to publish) independently of an author's own assessment of the value of that feedback.

We thus conclude that, from the point of view of epistemic consequentialism, there is presently little reason to believe that a loss of the epistemic sorting function of pre-publication peer review would be a loss to science. Inclusion in the literature does not do much to vouch for the quality of a paper; the evidence does not favor the hypothesis that reviewers are selecting for some latent epistemic quality in order to sort into appropriate journals; and the ideal underlying the claimed benefits of epistemic sorting is dubious. While peer reviewers do give potentially valuable feedback, there is no particular reason to think that changes in how scientists decide to spend their time would make things worse in this regard, and (per our arguments in section 3) some reason to think that they would make things better.

4.2 Malpractice Detection

The other way peer review might uphold epistemic standards is through malpractice detection. However, once again, the literature does not support this. A number of prominent cases of fraudulent research managed to sail through peer review. Upon investigation into the behavior of those involved it was found there was no reason to think that peer reviewers or editors were especially negligent in their duties (Grant 2002, 3). Peer reviewers report unwillingness to challenge something as fraudulent even where they have some suspicion that this is so, and avoid the charge (Francis 1989, 11–12). A criminologist who looked into fraudulent behavior in science reported that “virtually no fraudulent procedures have been detected by referees because reading a paper is neither a replication nor a lie-detecting device” (Ben-Yehuda 1986, 6). A more recent survey of the evidence found, at the least, no consistent pattern in journals’ self-reported ability to detect and weed out fraudulent results (Anderson et al. 2013, 235).

Even if the prospect of peer review puts some people off committing fraud, the fact that it is so unreliable at detecting fraud suggests that this is a very fragile deterrence system indeed. Even this psychological deterrence would be rapidly undermined by more adventurous souls, or those pushed by desperation, since many would quickly learn that pre-publication peer review is a paper tiger.

Conversely, there are various ways for malpractice detection to operate in the absence of peer review. These include motive modification (Nosek et al. 2012, Bright 2017a), encouraging post-publication replication and scrutiny (Bruner 2013, Romero 2017), and the sterner inculcation of the norms of science coupled with greater expectation of oversight among coworkers (Braxton 1990). All of these methods of deterring fraud or meliorating its effects would still be available under our proposal.

What evidence we now have gives little reason to suppose that abolishing pre-publication peer review is any great loss to malpractice detection. Thus in this regard our proposal would make no great difference to the epistemic health of science. Combining this with the discussion of epistemic sorting, we conclude there is presently no reason to believe pre-publication peer review is adding much value to science by upholding epistemic standards.

4.3 Herding Behavior

Where above we argued that pre-publication peer review is not making a positive difference often claimed for it, in this section we downplay a potential benefit of our proposal. A consistent worry about scientific behavior is that it is subject to fads or, in any case, some sort of undesirable herding behavior (see, e.g. Chargaff 1976, Abrahamson 2009, Strevens 2013). A natural thought is that pre-publication peer review encourages this, since by its nature it means that to get new ideas out there one must convince one's peers that the work is impressive and interesting. It has thus been claimed that pre-publication peer review encourages unambitious within-paradigm work that unduly limits the range of scientific activity (Francis 1989, 12). Reducing the incentive to herd might thus be claimed as a potential benefit of our proposal. However, we are not convinced that it is pre-publication peer review that is doing the harmful work here.

As mentioned above, our proposal eliminates or significantly reduces the importance of short-run credit, the credit that accrues to one in virtue of publishing in a (more or less prestigious) scientific journal. Long-run credit, on the other hand, is left untouched. Under any sort of credit system, a scientist needs to do work that the community will pay attention to, build upon, and recognize her for. The mere fact that (she believes that) her peers are interested in a topic and liable to respond to it is thus still positive reason to adopt a topic. This is true even if the scientist would not judge that topic to be the best use of her time if she were (hypothetically) free from the social pressures and constraints of the scientific credit system.

The best that could be said about our proposal in this regard is that scientists would not specifically have to pass a jury of peers before getting their work out there. But given that we anticipate continued competition for the attention of scientific coworkers, it is hard to say what the net effect in encouraging more experimental or less conformist scientific work would be.

Whatever conformist effects the credit incentive has (see also the discussion immediately below) do not depend on whether it is short- or long-run credit one seeks. The conformism comes from the fact that credit incentives focus scientists' attention on the predicted reaction of their fellow scientists to their work. Pre-publication peer review might make this fact especially salient by bringing manuscripts before a jury of peers before they may be entered into the literature. But even without pre-publication peer review the credit-seeking scientist must be focused on her peers' opinions. So there is no particular reason to think that removing the pre-publication scrutiny of manuscripts will free scientists from their own anticipations of the fads and fashions of their day.

4.4 Long-Run Credit

We end this section by noting that many of the effects of the credit economy of science studied by social epistemologists really concern long-run credit rather than the short-run credit affected by retaining or eliminating pre-publication peer review. This point is not restricted to herding behavior.

For instance, social epistemologists have studied both the incentive to collaborate, and various iniquities that can arise when scientists do not start with equal power when deciding who shall do what work and how they shall be credited (Harding 1995, Boyer-Kassem and Imbert 2015, Bruner and O'Connor 2017, O'Connor and Bruner forthcoming). Whether or not manuscripts would have to pass pre-publication peer review in order to enter the scientific literature, there would still be benefits in the long run to collaboration, and (alas) there would still be social inequalities that allow for iniquities to manifest in the scientific prestige hierarchy.

For another example, social epistemologists have studied the ways in which the credit incentive encourages different strategies for developing a research profile or molding one's scientific personality to be more or less risk-taking (Weisberg and Muldoon 2009, Alexander et al. 2015, Thoma 2015). Once again, pre-publication peer review plays no particular role in the analysis. The incentives to differentiate oneself from one's peers (without straying too far from the beaten path) and to mold one's personality accordingly exist independently of pre-publication peer review.

Two especially influential streams of work in the social epistemology of science have been the study of the division of cognitive labor (Kitcher 1990, Strevens 2003), and the role of credit in providing a spur to work in situations with a risk of under-production (Dasgupta and David 1994, Stephan 1996). These two streams have directed the focus of the field, and have formed some of the chief defenses of the credit economy of science as it now stands (but see Zollman 2018, for a more critical take).

We mention them here because pre-publication peer review or short-run credit again plays no particular role in the analyses offered by these papers. What drives their results is scientists' expectation that genuine scientific achievement will be recognized with credit. As we have argued above, it is long-run credit that best tracks genuine scientific achievement, and so it is long-run rather than short-run credit that grounds scientists' expectation in this regard. So in social epistemologists' most prominent defenses of the credit economy of science, long-run credit (while not named such) is the mechanism underlying the claimed epistemic benefits of the credit economy.

5 Difficulties For Our Proposal

We have discussed some benefits that would predictably accrue from abolishing peer review and some ways in which its apparent benefits are either under-evidenced or better attributed to the effects of long-run credit, which our proposal leaves untouched. We now discuss some cases which we take to be more problematic for our proposal—but by this point we hope to have at least convinced the reader that pre-publication peer review rests on shakier theoretical grounds than its widespread acceptance may lead one to suppose.

5.1 A Guarantee For Outsiders

One purpose pre-publication peer review serves is providing a guarantee to interested but non-expert parties. Science journalists, policy makers, scientists from outside the field the manuscript is aimed at, or interested non-scientists can take the fact that something has passed peer review as a stamp of approval from the field. At a minimum, peer review guarantees that outsiders are focusing on work that has convinced at least one relatively disinterested expert that the manuscript is worthy of public viewing. Given that there are real dangers to irresponsible science journalism or public action that is seen to be based on science that is not itself trustworthy (Bright 2018, §4), and that it is hard for non-experts to make the relevant judgment calls themselves, having a social mechanism to provide this kind of guarantee for outsiders is useful.

It is difficult to predict in advance what norms would come to exist for science journalists in the absence of pre-publication peer review. We thus first and foremost call for empirical research on this issue, possibly by studying what has happened in parts of mathematics and physics that already operate broadly along the lines we suggest (Gowers 2017).

However, against the presumption that things would be worse, we have two points to make. As the recent replication crisis has made clear, the value of peer review as a stamp of approval should not be overstated. There are reasons to doubt that peer review reliably succeeds in filtering out false results. We give three of them. First, peer reviewers face difficulties in actually assessing manuscripts—and just about anything can pass peer review eventually—as discussed under the heading of 'epistemic sorting' in section 4.1. Second, there are problems with the standards we presently use to evaluate manuscripts, in particular with the infamous threshold for statistical significance used in many fields (Ioannidis 2005, Benjamin et al. 2018). And third, deeper features of the incentive structure of science make replicability problems endemic (Smaldino and McElreath 2016, Heesen 2017c). Using peer review as a stamp of approval may just be generating expert overconfidence (Angner 2006), without the epistemic benefits of greater reliability that would back this confidence up.

For the second part of our reply, recall that it is only pre-publication peer review that we seek to eliminate. We do not object to post-publication peer review resulting in papers being selected for inclusion in journals which mark the community's approval of such work, ideally after due and broad-based evaluation. If some such system were implemented then outsiders could use inclusion in such a journal as their marker of whether work is soundly grounded in the relevant science.

If such a stamp of approval from a journal or other communally recognized institution only comes a number of months or years after something is first published then we would expect it to represent a more well-considered judgment. Note that this would not necessarily slow the diffusion of knowledge as under the present system the same paper would have spent time hidden from view going through pre-publication peer review. The end result might not even be all that different from what happens in the present system, except that post-publication peer review would take into account more of the response or uptake from the wider scientific community. Thus it would more closely approximate the considered judgment of the community, as ultimately reflected in the long-run credit accorded to the paper.

5.2 A Runaway Matthew Effect

The second problem we are less confident we can deal with is that of exacerbating the Matthew effect. This is the phenomenon, first identified by Merton (1968), of antecedently more famous authors being credited more for work done simultaneously or collaboratively, even if the relative size or skill of their contribution does not warrant a larger share of the reward. Arguably the present system helps put a damper on the Matthew effect, allowing a junior or less prestigious author to secure attention for their work by publishing in a high profile journal. Without such a mechanism to grab the attention of the field, perhaps scientists would just decide what to pay attention to based on their prior knowledge of the author or recommendation from others. This would strengthen the effects of networks of patronage and prestige bias favoring fancy universities. Thus squandering valuable opportunities to learn from those who were not initially lucky in securing a prestigious position or patronage from the already established.

While some have defended the Matthew effect (Strevens 2006), we will not go that route in defending our proposal for two reasons. First, the Matthew effect can perpetuate iniquities that themselves harm the generation and dissemination of knowledge (Bruner and O'Connor 2017). Second, even if it could be justified at the level of individual publications, its long-term effects are epistemically harmful. The scientific community allocates the resources necessary for future work on the basis of its recognition of past performance. So if there is excess reward for some and unfair passing over of others at the present stage of inquiry, this will ramify through to future rounds of inquiry, misallocating resources to people whose accomplishments do not fully justify their renown (Heesen 2017a). Hence on grounds of epistemic consequentialism we take seriously the problem of a runaway Matthew effect.

As mentioned, due to the pressures of credit-seeking and their own curiosity, scientists would still have incentive to read others' work and adapt it to suit their own projects. There is always a chance that valuable knowledge may be gathered from the work of one who has been ignored, which could provide an innovative edge. To some extent this creates opportunities for arbitrage: if the Matthew effect ever became especially severe there would be a credit incentive to specialize in seeking out the work of scientists who are not getting much attention. The lesson here is that the Matthew effect can only ever be so severe, before the credit incentive starts providing counter-veiling motivations.

However, this does not fully solve our problem. Moreover, so long as resource allocation is tied to recognition of past performance the differences in recognition generated by the Matthew effect can and often do become self-fulfilling prophesies, as those with more gain the resources to do better in the future, and those without are starved of the resources necessary to show their worth.

It is not clear where to go from here. From the above it may seem like a solution would be to pair our proposal with a call to loosen the connection between recognition of a scientist's greatness based on their past performance and resource allocation. Indeed, this may well be independently motivated (Avin forthcoming, Heesen 2017a, §6). However, even short of this far-reaching change, we feel at present that this matter deserves more study rather than any definitive course of action.

Our present thought is that this is a very speculative objection, and there is no empirical evidence to back up the claim that eliminating pre-publication peer review will have dire consequences in this regard. In particular, while the present system may (rarely) allow a relative outsider to make a big splash, the common accusation of prestige bias in peer review (Lee et al. 2013, 7) suggests that on the whole pre-publication peer review may contribute to the Matthew effect rather than curtailing it.

More specifically, the Matthew effect can be made worse by peer review when anonymity breaks down in ways that systematically favor antecedently famous scientists. If this gives famous scientists more opportunities to publish papers, then our system may provide welcome relief, since it allows more people to get their papers out there. Hence whether our proposal makes the Matthew effect worse or better depends on whether the stronger influence would be who gets into the conversation (for which pre-publication peer review can exacerbate the Matthew effect), or who gets listened to once the conversation has begun (for which our proposal looks more problematic). Presently we cannot say which is the more significant effect. So, while we grant that a runaway Matthew effect may occur under our system, we prefer to stress that at this point it is just not known whether the Matthew effect will be worse with or without pre-publication peer review.

What we propose is a large change, involving freeing up a lot of time and opening it up to more self-direction on the part of scientists, and it is not clear what sort of institutional changes it would be paired with. With more study of epistemic mechanisms designed especially to promote the work of junior or less prestigious scientists there might be found some way of surmounting the problem of a runaway Matthew effect, should it arise. Ultimately, only empirical evidence can settle these questions. Given the clear benefits and the unclear downsides of our proposal, we hope at minimum to inspire a more experimental attitude towards peer review.

6 Conclusion

Pre-publication peer review is an enormous sink of scientists' time, effort, and resources. Adopting the perspective of epistemic consequentialism and reviewing the literature on the philosophy, sociology, and social epistemology of science, we have argued that we can be confident that there would be benefits from eliminating this system, but have no strong reasons to think there will be disadvantages. There is hence a kind of weak dominance or Pareto argument in favor of our proposal.

To simplify things, imagine forming a decision matrix, with rows corresponding to 'Keeping pre-publication peer review' and 'Eliminating pre-publication peer review'. The columns would each be labeled with an issue studied by science scholars which we have surveyed here: gender bias in the literature, speed of dissemination of knowledge, efficient allocation of scientists' time and attention, etc. For each column, if there is a clear reason to think that either keeping or eliminating pre-publication scientific peer review does better according to the standards of epistemic consequentialism, place a 1 in the row of that option, and a 0 in the other. If there is no reason to favor either according to present evidence, put a 0 in both rows.

Our present argument could then be summarized with: as it stands, the only 1s in such a table would appear in the row for eliminating pre-publication peer review. We thus advocate eliminating pre-publication peer review. Journals could still exist as a forum for recognizing and promoting work that the community as a whole perceives as especially meritorious and wishes to recommend to outsiders. Scientists would still have every reason to read, respond to, and consider the work of their peers; pre-publication peer review is not the primary drive behind either the intellect's curiosity or the will's desire for recognition, and either of those suffice to motivate such behaviors.

The overall moral to be drawn mirrors that of our invocation of the importance of long-run over short-run credit. The best guarantor of the long run epistemic health of science is science: the organic engagement with each others' ideas and work that arises from scientists deciding for themselves how to allocate their cognitive labor, and doing the hard work of replicating and considering from new angles those ideas that have been opened up to the scrutiny of the community. All this would continue without pre-publication peer review, and the best you can say for the system that currently uses up so much of our time and resources is that it often fails to get in the way.

References

Eric Abrahamson. Necessary conditions for the study of fads and fashions in science. *Scandinavian Journal of Management*, 25(2):235–239, 2009. doi: 10.1016/j.scaman.2009.03.005. URL `http://dx.doi.org/10.1016/j.scaman.2009.03.005`.

Jason McKenzie Alexander, Johannes Himmelreich, and Christopher Thompson. Epistemic landscapes, optimal search, and the division of cognitive labor. *Philosophy of Science*, 82(3):424–453, 2015. doi: 10.1086/681766. URL http://dx.doi.org/10.1086/681766.

Melissa S. Anderson, Emily A. Ronning, Raymond De Vries, and Brian C. Martinson. Extending the Mertonian norms: Scientists' subscription to norms of research. *The Journal of Higher Education*, 81(3):366–393, 2010. ISSN 1538-4640. doi: 10.1353/jhe.0.0095. URL https://muse.jhu.edu/journals/journal_of_higher_education/v081/81.3.anderson.html.

Melissa S. Anderson, Marta A. Shaw, Nicholas H. Steneck, Erin Konkle, and Takehito Kamata. Research integrity and misconduct in the academic profession. In Michael B. Paulsen, editor, *Higher Education: Handbook of Theory and Research*, volume 28, chapter 5, pages 217–261. Springer, Dordrecht, 2013. doi: 10.1007/978-94-007-5836-0_5. URL http://dx.doi.org/10.1007/978-94-007-5836-0_5.

Erik Angner. Economists as experts: Overconfidence in theory and practice. *Journal of Economic Methodology*, 13(1):1–24, 2006. doi: 10.1080/13501780600566271. URL http://dx.doi.org/10.1080/13501780600566271.

Shahar Avin. Centralised funding and epistemic exploration. *The British Journal for the Philosophy of Science*, forthcoming. doi: 10.1093/bjps/axx059. URL http://dx.doi.org/10.1093/bjps/axx059.

Nachman Ben-Yehuda. Deviance in science: Towards the criminology of science. *British Journal of Criminology*, 26(1):1–27, 1986. doi: 10.1093/oxfordjournals.bjc.a047577. URL http://dx.doi.org/10.1093/oxfordjournals.bjc.a047577.

Daniel J. Benjamin, James O. Berger, Magnus Johannesson, Brian A. Nosek, E.-J. Wagenmakers, Richard Berk, Kenneth A. Bollen, Björn Brembs, Lawrence Brown, Colin Camerer, et al. Redefine statistical significance. *Nature Human Behaviour*, 2(1):6–10, 2018. ISSN 2397-3374. doi: 10.1038/s41562-017-0189-z. URL `http://dx.doi.org/10.1038/s41562-017-0189-z`.

Lutz Bornmann. Scientific peer review. *Annual Review of Information Science and Technology*, 45(1):197–245, 2011. ISSN 1550-8382. doi: 10.1002/aris.2011.1440450112. URL `http://dx.doi.org/10.1002/aris.2011.1440450112`.

Thomas Boyer. Is a bird in the hand worth two in the bush? Or, whether scientists should publish intermediate results. *Synthese*, 191(1):17–35, 2014. ISSN 0039-7857. doi: 10.1007/s11229-012-0242-4. URL `http://dx.doi.org/10.1007/s11229-012-0242-4`.

Thomas Boyer-Kassem and Cyrille Imbert. Scientific collaboration: Do two heads need to be more than twice better than one? *Philosophy of Science*, 82(4):667–688, 2015. ISSN 00318248. URL `http://www.jstor.org/stable/10.1086/682940`.

John M. Braxton. Deviance from the norms of science: A test of control theory. *Research in Higher Education*, 31(5):461–476, 1990. doi: 10.1007/BF00992713. URL `http://dx.doi.org/10.1007/BF00992713`.

Liam Kofi Bright. On fraud. *Philosophical Studies*, 174(2):291–310, 2017a. ISSN 1573-0883. doi: 10.1007/s11098-016-0682-7. URL `http://dx.doi.org/10.1007/s11098-016-0682-7`.

Liam Kofi Bright. Decision theoretic model of the productivity gap. *Erkenntnis*, 82(2):421–442, 2017b. ISSN 1572-8420. doi: 10.1007/s10670-016-9826-6. URL `http://dx.doi.org/10.1007/s10670-016-9826-6`.

Liam Kofi Bright. Du Bois' democratic defence of the value free ideal. *Synthese*, 195(5):2227–2245, 2018. ISSN 1573-0964.

doi: 10.1007/s11229-017-1333-z. URL http://dx.doi.org/10.1007/s11229-017-1333-z.

Liam Kofi Bright, Haixin Dang, and Remco Heesen. A role for judgment aggregation in coauthoring scientific papers. *Erkenntnis*, 83(2):231–252, 2018. ISSN 1572-8420. doi: 10.1007/s10670-017-9887-1. URL http://dx.doi.org/10.1007/s10670-017-9887-1.

Justin Bruner and Cailin O'Connor. Power, bargaining, and collaboration. In Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, editors, *Scientific Collaboration and Collective Knowledge*, chapter 7, pages 135–157. Oxford University Press, Oxford, 2017.

Justin P. Bruner. Policing epistemic communities. *Episteme*, 10(4):403–416, Dec 2013. ISSN 1750-0117. doi: 10.1017/epi.2013.34. URL http://dx.doi.org/10.1017/epi.2013.34.

Erwin Chargaff. Triviality in science: A brief meditation on fashions. *Perspectives in Biology and Medicine*, 19(3):324–333, 1976. doi: 10.1353/pbm.1976.0011. URL http://dx.doi.org/10.1353/pbm.1976.0011.

Diana Crane. The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *The American Sociologist*, 2(4):195–201, 1967. ISSN 00031232. URL http://www.jstor.org/stable/27701277.

Partha Dasgupta and Paul A. David. Toward a new economics of science. *Research Policy*, 23(5):487–521, 1994. ISSN 0048-7333. doi: 10.1016/0048-7333(94)01002-1. URL http://www.sciencedirect.com/science/article/pii/0048733394010021.

Margaret Eisenhart. The paradox of peer review: Admitting too much or allowing too little? *Research in Science Education*, 32(2):241–255, 2002. ISSN 1573-1898. doi: 10.1023/A:1016082229411. URL http://dx.doi.org/10.1023/A:1016082229411.

Edzard Ernst, T. Saradeth, and Karl Ludwig Resch. Drawbacks of peer review. *Nature*, 363(6427):296, 1993. doi: 10.1038/363296a0. URL `http://dx.doi.org/10.1038/363296a0`.

Henry Etzkowitz, Stefan Fuchs, Namrata Gupta, Carol Kemelgor, and Marina Ranga. The coming gender revolution in science. In Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, editors, *The Handbook of Science and Technology Studies*, chapter 17, pages 403–428. MIT Press, Cambridge, third edition, 2008. ISBN 9780262083645.

Daniele Fanelli. Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE*, 5(4):e10271, Apr 2010. doi: 10.1371/journal.pone.0010271. URL `http://dx.doi.org/10.1371/journal.pone.0010271`.

Jere R. Francis. The credibility and legitimation of science: A loss of faith in the scientific narrative. *Accountability in Research: Policies and Quality Assurance*, 1(1):5–22, 1989. doi: 10.1080/08989628908573770. URL `http://dx.doi.org/10.1080/08989628908573770`.

Alvin I. Goldman. *Knowledge in a Social World*. Oxford University Press, Oxford, 1999. ISBN 0198237774.

Timothy Gowers. The end of an error? *The Times Literary Supplement*, October 2017. URL `https://www.the-tls.co.uk/articles/public/the-end-of-an-error-peer-review/`. Editorial.

Paul M. Grant. Scientific credit and credibility. *Nature Materials*, 1:139–141, 2002. doi: 10.1038/nmat756. URL `http://dx.doi.org/10.1038/nmat756`.

Sandra Harding. "Strong objectivity": A response to the new objectivity question. *Synthese*, 104(3):331–349, 1995. doi: 10.1007/BF01064504. URL `http://dx.doi.org/10.1007/BF01064504`.

Remco Heesen. Academic superstars: Competent or lucky? *Synthese*, 194 (11):4499–4518, 2017a. ISSN 1573-0964. doi: 10.1007/s11229-016-1146-5. URL http://dx.doi.org/10.1007/s11229-016-1146-5.

Remco Heesen. Communism and the incentive to share in science. *Philosophy of Science*, 84(4):698–716, 2017b. ISSN 0031-8248. doi: 10.1086/693875. URL http://dx.doi.org/10.1086/693875.

Remco Heesen. Why the reward structure of science makes reproducibility problems inevitable. Manuscript, September 2017c. URL http://remcoheesen.files.wordpress.com/2015/03/rewards-and-reproducibility2.pdf.

Remco Heesen. When journal editors play favorites. *Philosophical Studies*, 175(4):831–858, 2018. ISSN 0031-8116. doi: 10.1007/s11098-017-0895-4. URL http://dx.doi.org/10.1007/s11098-017-0895-4.

Remco Heesen, Liam Kofi Bright, and Andrew Zucker. Vindicating methodological triangulation. *Synthese*, forthcoming. ISSN 1573-0964. doi: 10.1007/s11229-016-1294-7. URL http://dx.doi.org/10.1007/s11229-016-1294-7.

Erin Hengel. Publishing while female: Are women held to higher standards? Evidence from peer review. Manuscript, August 2018. URL http://www.erinhengel.com/research/publishing_female.pdf.

David L. Hull. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. The University of Chicago Press, Chicago, 1988. ISBN 0226360504.

John P. A. Ioannidis. Why most published research findings are false. *PLoS Medicine*, 2(8):e124, Aug 2005. doi: 10.1371/journal.pmed.0020124. URL http://dx.doi.org/10.1371/journal.pmed.0020124.

Saana Jukola. A social epistemological inquiry into biases in journal peer review. *Perspectives on Science*, 25(1):124–148, 2017. doi: 10.1162/POSC_a_00237. URL http://dx.doi.org/10.1162/POSC_a_00237.

J. Katzav and K. Vaesen. Pluralism and peer review in philosophy. *Philosophers' Imprint*, 17(19):1–20, 2017. URL http://hdl.handle.net/2027/spo.3521354.0017.019.

Philip Kitcher. The division of cognitive labor. *The Journal of Philosophy*, 87(1):5–22, 1990. ISSN 0022362X. URL http://www.jstor.org/stable/2026796.

Richard L. Kravitz, Peter Franks, Mitchell D. Feldman, Martha Gerrity, Cindy Byrne, and William M. Tierney. Editorial peer reviewers' recommendations at a general medical journal: are they reliable and do editors care? *PLoS ONE*, 5(4):e10072, 2010. doi: 10.1371/journal.pone.0010072. URL http://dx.doi.org/10.1371/journal.pone.0010072.

Bruno Latour and Steve Woolgar. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton, second edition, 1986.

Carole J. Lee. The limited effectiveness of prestige as an intervention on the health of medical journal publications. *Episteme*, 10(4):387–402, 2013. doi: 10.1017/epi.2013.35. URL http://dx.doi.org/10.1017/epi.2013.35.

Carole J. Lee. Revisiting current causes of women's underrepresentation in science. In Jennifer Saul and Michael Brownstein, editors, *Implicit Bias and Philosophy Volume 1: Metaphysics and Epistemology*, chapter 2.5, pages 265–282. Oxford University Press, Oxford, 2016. doi: 10.1093/acprof:oso/9780198713241.001.0001. URL http://dx.doi.org/10.1093/acprof:oso/9780198713241.001.0001.

Carole J. Lee, Cassidy R. Sugimoto, Guo Zhang, and Blaise Cronin. Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1):2–17, 2013. ISSN 1532-2890. doi: 10.1002/asi.22784. URL http://dx.doi.org/10.1002/asi.22784.

Christin List and Robert E. Goodin. Epistemic democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*, 9(3):277–306, 2001. ISSN 1467-9760. doi: 10.1111/1467-9760.00128. URL http://dx.doi.org/10.1111/1467-9760.00128.

Helen E. Longino. *Science as Social Knowledge*. Princeton University Press, 1990.

Karen Seashore Louis, Lisa M. Jones, and Eric G. Campbell. Macroscope: Sharing in science. *American Scientist*, 90(4):304–307, 2002. ISSN 00030996. URL http://www.jstor.org/stable/27857685.

Bruce Macfarlane and Ming Cheng. Communism, universalism and disinterestedness: Re-examining contemporary support among academics for Merton's scientific norms. *Journal of Academic Ethics*, 6(1):67–78, 2008. ISSN 1570-1727. doi: 10.1007/s10805-008-9055-y. URL http://dx.doi.org/10.1007/s10805-008-9055-y.

Robert K. Merton. A note on science and democracy. *Journal of Legal and Political Sociology*, 1(1–2):115–126, 1942. Reprinted in Merton (1973, chapter 13).

Robert K. Merton. Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22(6):635–659, 1957. ISSN 00031224. URL http://www.jstor.org/stable/2089193. Reprinted in Merton (1973, chapter 14).

Robert K. Merton. The Matthew effect in science. *Science*, 159(3810):56–63,

1968. ISSN 00368075. URL http://www.jstor.org/stable/1723414. Reprinted in Merton (1973, chapter 20).

Robert K. Merton. Behavior patterns of scientists. *The American Scholar*, 38 (2):197–225, 1969. ISSN 00030937. URL http://www.jstor.org/stable/41209646. Reprinted in Merton (1973, chapter 15).

Robert K. Merton. *The Sociology of Science: Theoretical and Empirical Investigations*. The University of Chicago Press, Chicago, 1973. ISBN 0226520919.

Brian A. Nosek, Jeffrey R. Spies, and Matt Motyl. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7(6):615–631, 2012. doi: 10.1177/1745691612459058. URL http://pps.sagepub.com/cgi/content/abstract/7/6/615.

Cailin O'Connor and Justin Bruner. Dynamics and diversity in epistemic communities. *Erkenntnis*, forthcoming. ISSN 1572-8420. doi: 10.1007/s10670-017-9950-y. URL http://dx.doi.org/10.1007/s10670-017-9950-y.

Slobodan Perović, Sandro Radovanović, Vlasta Sikimić, and Andrea Berber. Optimal research team composition: data envelopment analysis of Fermilab experiments. *Scientometrics*, 108(1):83–111, 2016. doi: 10.1007/s11192-016-1947-9. URL http://dx.doi.org/10.1007/s11192-016-1947-9.

Douglas P. Peters and Stephen J. Ceci. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, 5(2):187–195, 1982. doi: 10.1017/S0140525X00011213. URL http://dx.doi.org/10.1017/S0140525X00011213.

Katarina Prpić. Gender and productivity differentials in science. *Scientometrics*, 55(1):27–58, 2002. ISSN 0138-9130. doi: 10.1023/A:1016046819457. URL http://dx.doi.org/10.1023/A:1016046819457.

RIN. Activities, costs and funding flows in the scholarly communications system in the UK. Technical report, Cambridge Economic Policy Associates on behalf of the Research Information Network, 2008. URL http://rinarchive.jisc-collections.ac.uk/our-work/communicating-and-disseminating-research/activities-costs-and-funding-flows-scholarly-commu.

Felipe Romero. Novelty versus replicability: Virtues and vices in the reward system of science. *Philosophy of Science*, 84(5):1031–1043, 2017. ISSN 0031-8248. doi: 10.1086/694005. URL http://dx.doi.org/10.1086/694005.

Jennifer Saul. Implicit bias, stereotype threat, and women in philosophy. In Katrina Hutchison and Fiona Jenkins, editors, *Women in Philosophy: What Needs to Change?*, chapter 2, pages 39–60. Oxford University Press, Oxford, 2013.

Paul E. Smaldino and Richard McElreath. The natural selection of bad science. *Royal Society Open Science*, 3(9), 2016. doi: 10.1098/rsos.160384. URL http://rsos.royalsocietypublishing.org/content/3/9/160384.

Richard Smith. Peer review: a flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99(4):178–182, 2006. URL http://jrs.sagepub.com/content/99/4/178.short.

Paula E. Stephan. The economics of science. *Journal of Economic Literature*, 34(3):1199–1235, 1996. URL http://www.jstor.org/stable/2729500.

Michael Strevens. The role of the priority rule in science. *The Journal of Philosophy*, 100(2):55–79, 2003. ISSN 0022362X. URL http://www.jstor.org/stable/3655792.

Michael Strevens. The role of the Matthew effect in science. *Studies in History and Philosophy of Science Part A*, 37(2):159–170, 2006. ISSN 0039-3681. doi: http://dx.doi.org/10.1016/j.shpsa.2005.07.009. URL http://www.sciencedirect.com/science/article/pii/S0039368106000252.

Michael Strevens. Herding and the quest for credit. *Journal of Economic Methodology*, 20(1):19–34, 2013. doi: 10.1080/1350178X.2013.774849. URL http://dx.doi.org/10.1080/1350178X.2013.774849.

Michael Strevens. Scientific sharing: Communism and the social contract. In Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, editors, *Scientific Collaboration and Collective Knowledge*, chapter 1. Oxford University Press, Oxford, 2017. URL https://philpapers.org/rec/STRSSC-2.

Johanna Thoma. The epistemic division of labor revisited. *Philosophy of Science*, 82(3):454–472, 2015. ISSN 00318248. URL http://www.jstor.org/stable/10.1086/681768.

Virginia Valian. *Why So Slow? The Advancement of Women*. MIT Press, Cambridge, 1999. ISBN 9780262720311.

Richard Van Noorden. The true cost of science publishing. *Nature*, 495 (7442):426–429, 2013. ISSN 0028-0836. doi: 10.1038/495426a. URL http://dx.doi.org/10.1038/495426a.

Michael Weisberg and Ryan Muldoon. Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, 76(2):225–252, 2009. ISSN 00318248. URL http://www.jstor.org/stable/10.1086/644786.

Kevin J. S. Zollman. Optimal publishing strategies. *Episteme*, 6(2):185–199, Jun 2009. ISSN 1750-0117. doi: 10.3366/E174236000900063X. URL `http://dx.doi.org/10.3366/E174236000900063X`.

Kevin J. S. Zollman. The credit economy and the economic rationality of science. *The Journal of Philosophy*, 115(1):5–33, 2018. doi: 10.5840/jphil201811511. URL `http://dx.doi.org/10.5840/jphil201811511`.

Harriet Zuckerman and Jonathan R. Cole. Women in American science. *Minerva*, 13(1):82–102, 1975. ISSN 1573-1871. doi: 10.1007/BF01096243. URL `http://dx.doi.org/10.1007/BF01096243`.