Enslin, P. and Hedge, N. (2018) On peer review as the 'gold standard' in measuring research excellence: from secrecy to openness? *Journal of Philosophy of Education*, 52(3), pp. 379-396.

This is the peer reviewed version of the following article: Enslin, P. and Hedge, N. (2018) On peer review as the 'gold standard' in measuring research excellence: from secrecy to openness? *Journal of Philosophy of Education*, 52(3), pp. 379-396, which has been published in final form at http://dx.doi.org/10.1111/1467-9752.12312

http://eprints.gla.ac.uk/171817/

Deposited on: 23 October 2018

On peer review as the 'gold standard' in measuring research excellence: from secrecy to openness?

PENNY ENSLIN AND NICKI HEDGE[1]

ABSTRACT

As universities in the United Kingdom gear themselves up for the next Research Excellence Framework, REF2021, with peer review at its core, we critically re-visit the idea of peer review as a gold standard proxy for research excellence. We question the premise that anonymous peer review is a necessary and enabling condition for impartial, expert judgement. We argue that the intentions and supposed benefits underlying peer review and its associated concepts have become congealed in received discourse about research quality. Hence we explore the key conceptual issues raised by the nested assumptions and concepts that come into play in peer review as currently practised: primarily those of secrecy, anonymity, legitimacy, trust, impartiality, and openness. After delineating the benefits attributed to peer review, we contrast its declared virtues with its problematic features. We locate peer review in an audit culture in which the reviewer is an academic labourer. Drawing on recent trends in moral and political philosophy, we question the usefulness of the ideal of impartiality when tied to secrecy. Then we raise more deliberative, intersubjective possibilities for a revised understanding of peer review in the context of an academic community. Finally, we suggest ways in which the academic community could pursue quality in research by recasting peer review to be less secret and more open.

---

[1] penny.enslin@glasgow.ac.uk, nicki.hedge@glasgow.ac.uk

1 INTRODUCTION

> A path, a riddle, a jewel, an oath – anything can be secret so long as it is kept intentionally hidden, set apart in the mind of its keeper as requiring concealment. Sisela Bok (1989, p. 5)

Peer review, still regarded as the 'gold standard' of quality judgment in research publications, is a high stakes activity that can make or break academic careers. Writing as both reviewers and reviewed, we have benefitted from peer review that has entailed constructive critique enabling us to improve our papers. We ourselves review in that spirit, and acknowledge that author anonymity can provide rejected authors with a form of protective privacy, arguably a form of confidentiality designed to protect their dignity. But, inevitably too, we have received reviews that are hard to fathom; reviews that, in Bok's terms, are from supposed 'peers' whose identities are intentionally hidden, in words which read as riddles that sometimes allude to paths concealed. Following Bok (1989, p.9), our initial approach to secrecy is neutral and we acknowledge that 'a degree of concealment or openness accompanies all that human beings do or say'. With secrecy appropriate and necessary in some contexts, our focus is on secrets as 'intentionally hidden … requiring concealment' (Bok, 1989, p. 5). Hence we interrogate the intentions underpinning the secrecy of peer review, asking about its extent and consequences and questioning if concealment is actually required. Our scrutiny does not indicate a complete rejection of peer review. Rather we critically re-visit the idea of anonymised peer review as a gold standard proxy for research excellence questioning the assumption that it is a necessary and enabling condition for impartial, expert judgment.

With respect to publications[1], our scrutiny is timeous—as universities in the UK gear themselves up for the next Research Excellence Framework[2] (REF), with peer review at its core. Indeed, it was that 'gearing up', aptly described as an 'audit frenzy' to assess each university's 'REF readiness' (Lucas, 2017, p. 213), which initially motivated this work. While peer review, rather than bibliometrics, remains at the heart of the UK's REF, across Europe performance based research funding systems similarly and frequently inform institutional funding[3]. Beyond Europe, Australia, Hong Kong, New Zealand and South Africa deploy research performance based funding systems (Hicks,

2012) and so philosophical attention to peer review merits attention as an internationally practised process of ever increasing relevance for the academic labourer and for the quality of our work.

REF2021 throws peer review into sharp relief, casting a bright spotlight both on what is 'secret' and what intentions motivate that secrecy. As we examine why and how concealment imbues peer review, we explore the relationships between concealment and the exercise of power. In considering understandings and modes of peer review and its importance in today's academy, we also take heed of Biagioli's (2002, p. 7) view that in light of 'the remarkable epistemological and symbolic burden placed on peer review', too little research has analyzed it. Acknowledging an upsurge of empirical research and debate on peer review in the last decade, we note that much of that activity has occurred in the sciences, with philosophical examination remaining scant. While peer review is often discussed and is frequently the target of criticism, academics including philosophers of education still:

> … do not frame it as an intellectual subject. Instead, they either confine it to
> private conversations or treat it as one of the practical aspects of the
> profession. (Biagiolo, 2002, p.7)

We shift such conversations to a philosophical interrogation, arguing that the intentions and supposed benefits underlying peer review and its associated concepts have become congealed in received discourse about research quality. With the consequences of peer review increasingly significant for academic lives, we explore the key conceptual issues raised by the nested assumptions and concepts that come into play in peer review as currently practised: primarily those of *secrecy, anonymity, legitimacy, trust, impartiality,* and *openness*. In section 2, after delineating the benefits attributed to peer review and its importance in publishing and in the REF, we contrast its declared virtues with its problematic features, locating these in an audit culture in which the reviewer is also an academic labourer working in an environment of managerialism and systemic *distrust*. Section 3 then draws on recent trends in moral and political philosophy, to probe the principle of impartiality, questioning its usefulness when tied to secrecy and raising more deliberatively intersubjective possibilities for a revised understanding of peer review in the context of an academic community. In concluding, we move in section 4 towards recasting peer review in a more open mode, in ways that might better serve the interests and contributions of our work, with reference to alternatives and

enhancements to current practices. In doing so we suggest ways in which peer review might become more open through deliberative processes which include development for peer reviewers and ongoing debate in academic communities about what makes for quality research.

In suggesting ways of ameliorating the most pernicious aspects of secret peer review, we build on Conroy and Smith's (2017) recent arguments in this journal on the ethics of the REF and its effects on the higher education sector. In many respects this paper is a response to their call to 'go on speaking, not without a little irony' as we 'lean against the prevailing winds, for the deeper and more enduring ethical purposes and values of the university' (Conroy and Smith, 2017, p.14). Those prevailing winds have rendered peer view increasingly important and with far-reaching consequences that go beyond the quality of journals and the reputation of any academic discipline or university.

2 THE PEER REVIEWER AS ACADEMIC LABOURER

Pointing to the centrality of peer review in academic publishing, Jackson et al (2018, p.95) observe that it 'is seldom acknowledged publicly as a normal part of academic work'. Focussed on that academic work with respect to forms of anonymised pre- and post-publication peer review, we provide an account of understandings of peer review and its importance. We locate the peer reviewer, both publications reviewer and REF-related internal, institutional reviewer, as an academic labourer[4]—with that labour seldom acknowledged and, just as importantly, seldom theorised, publicly or institutionally. The very practice of peer review, in its various forms, is often defended as the best form of judgment we have on the grounds of impartiality and the maintenance of quality standards. But these defences merit scrutiny.

Attempts to define peer review frequently refer to 'expert' judgments of reviewers and maintaining standards for ensuring high quality research. Wilsdon et al state:

> Peer review is a general umbrella term for a host of expert-based review
> practices including the review of journal manuscripts, peer review of
> applications for funding and career promotions, and national peer review-
> based research assessments such as the REF. (2015, p. 59)

Acknowledging the variety of its meanings and forms, peer review is regarded by Wouters et al (2015, p. 44) as, 'without doubt the most important method of quality control in the sciences, the social sciences, arts and the humanities'. Similarly, the Stern Review (UK Government, 2014), with its very significant influence on REF2021, noted that academics had pointed to the import and advantages of peer review over a metrics-based approach to assessments of quality in publications[5]. However, Stern's report also acknowledges challenges in applying the 'gold standard of peer review' remarking that: 'At best, peer review is not a perfect "measure" ' (UK Government, 2014, p.4). But in an era of limited journal space and an unrelenting drive to 'publish or perish', peer review has long been seen as appropriate 'to referee the competition' in journals with limited space—although, with peers who may be researchers' competitors, 'prejudice and jealousy may lead to rejection and violation of academic freedom' (Berry, 1980, p. 639). Of course, types and degrees of secrecy in peer review vary. High ranking journals in the social sciences[6] frequently deploy a double blind peer review process in which the identities of the reviewed and reviewer are intentionally concealed[7]. We might more accurately refer to this form of review as 'double blind double' as it also entails that the reviewers and their reviews are concealed from each other[8], while journal editors know the identities of both author and reviewers[9].

In common with most academics, we regularly assume several roles within this review regime, most frequently as authors with our own work put forward for pre-publication, anonymous double blind peer review. Perhaps most usefully, however, we benefit from peer feedback in open discussion at conferences and with peers. By contrast, as reviewers, we regularly collude with secrecy by reviewing anonymously for specialist journals in philosophy of education as well as more generalist and interdisciplinary education journals requiring the same double blind review systems.

Across HE institutions, however, the role of reviewer has expanded to embrace reviewing internally because, using data from such peer review, universities assess their REF2021 readiness and implement strategies to enhance their chances of faring well in the next exercise. Institutional REF preparation across the UK continues apace as universities assess the readiness of individual academics to submit, almost always, at least one output predicted at 3* and preferably 4*, with a high quality submission carrying huge institutional advantages, both financial and reputational[10]. This

preparation also involves calibration exercises, in which internal reviewers are themselves reviewed to assess the accuracy of their assessments of their peers' work, frequently checked by external reviewers, with the expertise of members of previous REF panels much in demand. In REF2021 all staff with a research or research and teaching contract must be submitted with a minimum of one and maximum of five 'outputs', with each output assessed according to its *originality, significance,* and *rigour*[11]. While this paper is not exclusively REF focussed, the REF amplifies the urgency of scrutinising the peer review principle on which it depends. Following recent draft guidance on REF submission[12], Kernohan (2018) confirms that, 'at every stage it is human judgment rather than metrics that forms the backbone of the REF'. That human judgment will be fallible is a truism, but the rejoinder that the best processes of peer review have in-built safeguards to ameliorate such fallibility deserves scrutiny with particular respect to the reproduction of secrecy.

We are aware that practices currently underway across the sector in anticipation of REF2021 vary. Some internal, institutional REF related peer review activity is not anonymised at all, while some deploys a single blind process in which the authors' identities are not concealed from reviewers while reviewer identities remain secret to the author. As in the REF itself, this is a form of post-publication review. However, in the REF the single blind process, in which the identity of authors is revealed to panellists, is slightly modified with the identities of panel members collectively known to the public, while individual reviews and scores remain secret[13]. Additionally, while the criteria for 'excellence' are not concealed, peer review renders the processes of selection secret for entry to the REF in some institutions.

As we peer review, subject ourselves to peer review, and are reviewed as reviewers, there is, as we noted in our introduction, much at stake. The reviewer, reviewee and reviewed reviewer may have secret identities in some review processes but they are also selves whose identities are located in a discipline (in our case philosophy of education), and as employees in a higher education sector that is performatively driven and highly competitive. But peer review has tentacles beyond the REF: it is a high stakes activity for individual academics as well as their employer institutions. A requisite number of papers judged through several layers of peer review to be of an adequately high standard ensure one is 'REF-able'. But, following McCulloch (2017, pp. 512-3), such

judgments also enable academics to 'be employable and promotable, and, ironically, to gain access to the time and support necessary to facilitate the production of good quality research … not only in order to progress in their career, but also to keep their current job and avoid sanctions'. More positively, a good peer review and resulting publications can give one's work an audience, new networks of correspondents and potential collaborators, as well as enhanced influence in the field of study.

Of course, the secrecy underpinning such review processes has honourable intentions. While noting that 'the exact beginnings of anonymous peer review are a bit more vague than those of peer review itself', Gould (2010, p. 437) suggests anonymous peer review was originally intended 'to generate more candid evaluations unaffected by personal feelings or institutional biases'. Hence concealing the identity of reviewers is defended on the grounds that this offers protection from bias since reviewers are "third party" participants (Smith, 2006, p. 178). However, concludes Smith (2006, p. 179), editor of the *British Medical Journal* for 13 years:

> … we have little evidence on the effectiveness of peer review, but we have considerable evidence on its defects. In addition to being poor at detecting gross defects and almost useless for detecting fraud it is slow, expensive, profligate of academic time, highly subjective, something of a lottery, prone to bias, and easily abused.

Highlighting the subjective and secretive nature of much peer review, Hirsch et al (2017, e5) call for journal editors to ensure impartiality to protect against reviewer bias. We acknowledge the power of editors and editorial boards, and the extent to which academic communities rely on them to exercise fair and informed judgment, especially when some reviewers, on close scrutiny, can be biased, uninformed but nonetheless ready to make influential judgments of work with sometimes unfairly destructive and wide-ranging consequences for their authors. We will return to the responsibility of editors and interrogate the notion of impartiality as a defence for secrecy in the next section, but we turn now to the role of today's academic labourer acted upon and active in peer review.

In today's academy, the role of peer reviewer has extended from its initial construal as an apparently impartial judge of quality in journal publication, acting on behalf of a discipline and providing free labour to publishers[14], to that of employee reviewee and

reviewer. The peer reviewer today is hence a functionary increasingly complicit in the exercise of power, an academic labourer enabling her employer to manage fellow academics. We do not intend this observation to dismiss peer review as such, and as loyal employees of our own university and members of a disciplinary community of philosophers of education we do our best to contribute with integrity to all the types of review described here. But at this stage in our argument our point is to warn of the dangers of naively failing to recognise the problematic tensions now imbricated in the role of peer reviewer. So as we interrogate the role of the 'peer' and of secrecy in reviewing practices, it is necessary to recognize that confidentiality in peer review is one of several ways in which power operates secretly to normalize academics' professional behaviour (Baez, 2002, p. 168).

In exercising philosophical vigilance in the face of the appropriation of peer review into the 'audit explosion' that many believe poses considerable risk to the moral integrity of the university, the concept of trust has an acute relevance to our argument in two senses: implied distrust and unqualified trust. The first of these requires us to note, as Power has argued (1997), that although checking up is sometimes justified, the impulse to audit implies a degree of systemic distrust, a *need* to check up. While providing an account of ourselves 'sustains the fabric of normal human existence' (Power, 1997, p.1), checking up on one another through the many forms of monitoring that now characterise the 'audit society' comprises far more than a set of merely technical tasks. It shapes our very understandings of the issues it is supposed to address and constitutes a style of regulation with profound implications. 'Auditing may be a collection of tests and an evidence gathering task, but it is also a system of values and goals which are inscribed in the official programmes which demand it' (Power, 1997, p.7).

Whether external or internal to an organisation like a university, audit is never neutral. It has unintended cultural effects that alter, in the context we are discussing here, the peer and author selves as well as the institutions whose performance the audit intends to regulate. Yet, with its 'incentive effects', more auditing does not necessarily lead to greater accountability and the consequences as well as the intentions underpinning auditing's secrecy demand attention. Power's observation that 'Assumptions of distrust sustaining audit processes may be self-fulfilling as auditees adapt their behaviour strategically in response to the audit process, thereby becoming less trustworthy' (1997,

p. 135) has been vindicated as universities have, understandably, tried to 'game' the rules of each iteration of audits of research excellence over several decades. While the past practice of 'buying in' researchers offering a brace of probable 4* outputs has been headed off for REF2021 by a change in the rules, authors and reviewers are nonetheless closely attuned to the criteria that demand that outputs demonstrate significance, originality and rigour[15]. They will, to varying degrees and sometimes slavishly, adjust their practice to the culture of review and to writing to the rules of the coming audit of their research excellence. So with peer review a high stakes activity for all academics, it merits interrogation—not least with respect to its fallibility and the need to question the intentions and consequences of its secrecy in the current audit climate. This is particularly urgent if we, as an academic community, are to avoid peer review further infiltrating funding and appointment systems 'based on a fetishised image of this concept' (Derricourt, 2012, p. 137), in which, following secretive peer review, the identities and reputations of academic labourers may be shaped.

As the impulse to review all types of performance grows, so the pool of qualified peers is likely to become too small for the many acts of review required. With so much reviewing in demand and the expertise of reviewers as well as their time stretched and diluted across a greater volume of reviewing, the risks of blind review will increase. Peer review is already controversial. While conceding that it may sometimes be valuable, Butchard et al remark (2017, Introduction) that anecdotally there is plenty of evidence of peer review's pitfalls:

> Asked to share her own 'horror stories' in peer review for the *Times Higher Education*, Susan Bassnett[16] comments that 'it seems like a fine idea for work submitted to a journal, publisher or funding body to be assessed anonymously by independent experts', but fears peer review 'has grown into a monster' as a result of an increasing volume of work requiring review, with detrimental effects for both authors and reviewers.

In this vein, most authors of academic papers submitted to journals in the hope of publication are able to contribute to a collective narrative in which our experiences range from illuminating, constructive peer reviews that may result in revision enhancements and ultimate publication, to grumpy dismissal with minimal feedback or doctrinaire refusal to read a paper on its own merits. Fair and sensible decisions by editorial teams can ameliorate the worst of such tendencies, with experienced and

principled editors exercising careful judgment in managing and acting on peer reviews (Derricourt, 2012).

When peer review works well, the feedback it provides can be particularly instructive for early career researchers trying to learn the often opaque rules of academic practice. But while some reviewers conduct themselves as peers, others use the secrecy afforded by anonymous blind review to decline to act like peers who are equal participants in a shared academic community. Authors sometimes discover that blind reviewers seem to refer to criteria or expectations that were effectively secret at the time of submission. We are struck by the contingency inherent in processes of anonymous peer review as routinely practised. Sometimes we are lucky to have our work reviewed by a genuine peer who knows the field and bases her review not on what her own specific ideological stance may demand: she eschews judging the article on the basis of the paper she herself would have written or the different books she has read. Because luck and risk are part of the publication game, 'unqualified trust' is the second sense in which we deliberate on trust. Unqualified trust in the legitimacy of peer review is not warranted, granting peer review greater authority than it deserves. Under such conditions of secrecy, why might we trust those who have the power to make decisions to make those decisions wisely, especially if the decisions they make are sometimes so surprising and of such consequence for those whose work is being judged? In the courts, by contrast, unless circumstances are exceptional, justice must be seen to be done; the identities of witnesses and the jury of one's peers are not kept secret, and judgments as well as their rationales are publically available.

A possible rejoinder is that peer reviews are judgments that must occasionally take the form of gate keeping, on the assumption that some mistakes might be made, with some reviews perhaps too harsh but, if the bar is kept high, overall quality will ensue. In general, so this response might go, there is a good chance that those outputs that get past vigilant reviewers are probably of a sufficiently high quality to be worth publication. Occasional over-harsh reviews will at least support this worthy goal. The problem with this stance is that it confirms the contingency of peer review, that it is a crude measure that occasionally lets work later revealed to be of dubious merit through the gate, and fails to take seriously the collateral damage to the work and careers of those whose work is severely judged. So, too, this contingency may result in the non-

publication of papers that are later revealed to be original, significant and rigorous[17]. This discourages trust in a system whose procedures and underlying assumptions are of such immense consequence for all academic labourers.

Against this background, we might also ask why philosophers of education, for all their evident interest in critiquing performativity in educational institutions—as well as their close attention to analysing and defending concepts like professional agency and criticality—submit so meekly to the REF regime and participate in peer review in all of its forms and in all of their identities. We might usefully question our possible complicity in the more managerialist uses of peer review in the economy of the university. Of all researchers, philosophers of education, located as they are in Unit of Assessment 23 *Education,* might be expected to be alert to the language and ethics of the REF process and its requirements as well as to its avowed benefits for the academic community. Philosophers of education, so well-versed in theories of power and discourse, are located well to scrutinise the authority of peer review as a disciplinary mechanism. They are ideally placed to put the concepts we frequently use in critical comment about schooling or higher education to work in interrogating practices close to our own work and being as academics, including describing our publications in industrial terms as 'outputs', as if produced by 'knowledge workers' (Conroy and Smith, 2017, p. 10).

3 THE PEER REVIEWER AS JUDGE OF RESEARCH QUALITY: FROM 'IMPARTIALITY' TO PUBLIC DELIBERATION

The peer reviewer is routinely portrayed as an impartial judge of research quality, and it is widely held that, 'Impartial peer review is central to the scientific endeavour that all of us are engaged in…' (Miller et al, 2013, p.120). In its explanation of peer review the publisher Elsevier also emphasises the principle of impartiality, confidently asserting that 'Reviewer anonymity allows for impartial decisions – the reviewers will not be influenced by the authors'[18]. We question this widely held depiction of the peer reviewer as an impartial, knowledgeable judge of research quality. The very idea - indeed the possibility - of impartiality demands close scrutiny, as do the forms of

secrecy that inform the impartiality attributed to peer review. The assumption that anonymity fosters impartiality does not stand up to scrutiny.

Being an impartial reviewer implies that all submissions merit equal consideration, applying an appropriate set of criteria even-handedly. Impartiality fostered by anonymity thus precludes both favouring one's friends' work and taking a biased approach to those one dislikes or does not know, which might include those critical of one's own publications or who take a different ideological stance. Anonymity seems likely to head off the danger that a reviewer will warmly recommend work submitted by those known to her and perhaps those from her own university, aware of its competition with others to excel in the REF and to climb the league tables that feed off it[19]. Yet anonymity will not, in practice, guarantee this and it does not necessarily enhance impartiality. It is no protection when an anonymous author's arguments and theoretical assumptions - which can hardly be kept secret - offend a reviewer: for example, when a liberal feminist author errs, in the eyes of a 'poststructuralist' reviewer, by referring to young people who have been sexually exploited as 'victims'[20]. Feminist authors who offend malestream epistemologists by writing about women's knowledge, or philosophers of education writing in a particular tradition and having their work reviewed inflexibly by orthodox proponents of a rival tradition, can have that work dismissed from behind a screen of supposedly impartial anonymity, even if their named identities remain a secret. One reviewer's *bête noir* can seem perfectly reasonable to a reviewer operating in a different discipline or intellectual tradition: for example, a reviewer's mistaking an illustrative vignette for an empirical claim, or misinterpreting a conceptual distinction as committing the error of creating a 'binary', or confusing a paper that is 'well-written' with one that is philosophically rigorous – types of ignorance to which philosophers are particularly vulnerable at the hands of those from other fields. Philosophers of education may be also be vulnerable to internal review by colleagues judging their work not to be 'proper research' as it has neither collected nor analysed first or second order data.

Revisiting the principle of impartiality in academic publishing and its unquestioned status in the quality discourse of the university can benefit from attention to recent reconsideration of the ethics of impartiality in moral philosophy, which has been driven largely though not exclusively by feminist scholarship (e.g. Young, 1990). This work is

salient to both the ethics of peer review and the wider moral economy of the university. Friedman has acknowledged the allure of impartiality, suggesting that as a principle 'moral impartiality calls upon the moral agent to give equal consideration to the morally relevant interests of all persons' (1993, p. 3). But not only is partiality sometimes appropriate, as we will argue later; for Friedman 'the impartial standpoint is humanly impossible' (p. 3) and she recommends replacing pursuit of impartiality with paying attention to eliminating 'specific biases' from moral thought (p. 11). If Friedman is right in defending the identification and correction of specific biases as an indirect route towards the problematic ideal of impartiality, she might be interpreted as endorsing the now popular online unconscious or implicit bias training[21] that peer reviewers are commonly required to take before being let loose on others' 'outputs'. But her argument goes further:

> Coupled with the empirical assumption that people in general lack privileged access to their own biases, this approach, I suggest, invites us to appreciate the crucial role played by moral dialogue in the recognition of bias and, in this respect, the practical intersubjectivity of the enterprise of critical moral thinking.
> Friedman (1993, p. 11)

In her account of how others' biases could be corrected through the expression of competing points of view, Friedman (1993) draws on discourse ethics to contrast her defence of interpersonal dialogue in normative reflection with what Habermas (1979) has described as the 'monological' approach that assumes an isolated reasoner. Suggesting that monological theories 'overlook the possibilities for achieving impartiality that inhere in interpersonal dialogue', Friedman (1993, p.17) argues that a dialogic approach will enable people to 'together self-consciously seek to eliminate the biases in each other's moral thinking', arguing that authentic dialogue is a prerequisite 'for eliminating or minimizing bias in normative thinking'.

That our own biases are likely to remain concealed if we are forced to reflect unaided and in isolation is surely a reason to turn instead to interpersonal dialogue in reading and thinking about the research of others in peer review. Quality judgment here is as likely to benefit from such dialogue as is moral thinking. Peer review can fail to meet its best potential not only through bias, but also, occasionally, through sheer ignorance that could be corrected in communication with others. Serial, discrete and supposedly detached judgments of isolated anonymous reviewers, insulated from the opinions of

other peer reviewers, will not by definition be superior to those reached through more open, dialogical processes of review. Many experienced authors are able to cite examples of occasional breathtaking ignorance and carelessness revealed by reviewers whose lapses in expertise, confidently asserted from a position of anonymity, could be easily corrected if communication were possible. Problems of this kind point to the role of editors to occasionally set aside bad peer reviews which, if advice were followed, could make a paper worse rather than better, and to mediate between sometimes conflicting reviewer opinions - when dialogue between them could offer a far more rational review process. Interpersonal, open, even public, dialogue, making one's own thinking explicit and adjusting it in response to previously unnoticed considerations, offers a more reliable means to making informed and hence legitimate judgments of quality. Indeed, in her study of the concept of secrecy, Bok observes that having to argue publicly is likely to create 'the necessity to articulate one's position carefully, to defend it against unexpected counter arguments, to take opposing points of view into consideration, to reveal the steps of reasoning one has used, and to state openly the principles to which one appeals' (1989, p.114).

The imperfections to which peer review and its uses are prone do not lead us to reject it. Nor are we entirely opposed to the principle of impartiality. Instead, we turn to recent developments in democratic theory for a means to further recast our understanding of judgments of quality that avoid bias, through interpersonal deliberation. In doing so we work off Shatz's observation, in comparing peer review with democracy, that 'it is the worst form of evaluation - except for all the others' (Shatz, 2004, quoted in Butchard et al 2017, p. 18)—reflecting on the deliberative turn in democratic theory. Just as Churchill's famous observation is not a rejection of democracy, so too we do not aim to reject peer review so much as to critically evaluate its strengths, weaknesses and possibilities for reformulation, by now considering a way forward drawn from theories of democracy conceived as public reason.

As a response to the criticisms of democracy construed primarily as aggregation of the individual preferences of voters by counting their votes in elections or referenda, a form of decision-making that might be compared with measuring the quality of publications through metrics, the deliberative ideal is commonly described as embodying 'the "will of the people" formed through the public reasoning of its citizens' (Bohman, 1998, p.

14

401). In defending public deliberation among citizens who are free and equal, deliberative theory offers a path towards legitimate decision making for the common good (ibid.), even and especially in addressing complex problems in contexts of deep disagreement. The ideal of public reason offers a means to improved decision making. Deliberation that is free and open is more likely to be genuinely impartial, as the interests of all may be fairly taken into account in discourse that is open and dialogical.

The idea of 'deliberative impartiality' offers a legitimate process of collaborative decision-making instead of 'the ideal of the impartial transcendental subject' (Button and Garrett, 2016, p. 49). When decisions are reached through processes that involve interpersonal forms of reasoning and are public by way of offering, weighing and even rejecting reasons available to all affected, legitimacy is made possible through what Dryzek (2001) has called 'reflective assent'. Legitimacy can be enhanced by deliberation, in public decisions as well as in reaching decisions *about* publications, when there are opportunities for preferences to be transformed by confronting the preferences of others (Benhabib, 1996, p. 71). So, too, deliberative procedures generate legitimacy in part by imparting information, which no single individual can possess in full. Nor can one individual foresee all possible critical perspectives on a single issue (ibid). These deliberative ideals are envisaged as applicable not only to political institutions like parliaments and mass assemblies, but have been proffered in reflection on the workings of a range of institutions, 'a *plurality of modes of association* in which all affected can have the right to articulate their point of view' (Benhabib, 1996, p. 73). These overlapping and interlinked associations are seen by Benhabib to include voluntary associations and social movements as well as political parties. For us, this public sphere also embraces communities of academics and researchers.

Blind, anonymous peer review, conducted in secret, casts the peer reviewer as a discrete actor, insulated from the contamination of judgments reached by other reviewers. We contrast this problematic construal with the peer reviewer as member of a public, a disciplinary community with shared interests in voicing their views about the discipline and the literature that represents it in the wider field of educational research and the general public at large. A disciplinary community, with its mutual commitment to influencing the course of educational policy and practice, and educational research in other disciplines, is more likely to be collectively influential through interpersonal

deliberation than entrusting judgments about potential publications so strongly to isolated, anonymous reviewers. The community has an interest in exercising partiality in a collective commitment to support fellow scholars and for the health and survival of the discipline, rather than to impartiality simplistically conceived. Its members pay a price for the secrecy enforced by delegating judgments of quality to discrete individuals whose identities and opinions are concealed on the suspect assumption that this will make them impartial.

4 TOWARDS OPEN REVIEW

Admirable intentions do not necessarily produce their intended consequences. Our scrutiny of peer review has questioned the assumptions that underpin entrenched peer review practices that are now increasingly harnessed to an audit regime in which academic labourers are complicit in the exercise of power over one another. Our analysis of these assumptions, and the associated concepts that legitimize peer review as currently practised, has contrasted secrecy with openness, with close scrutiny of the concept of 'impartiality'. While not dismissing the principle of impartiality, we have argued instead for interpersonal dialogue through more open forms of deliberation about research quality, which we claim are more likely to foster legitimacy and trust, as well as enhancing judgments of quality. What changes in current peer review practices might this suggest?

Addressing 'inherent issues in what has been the gold standard of double-blind peer review', Open Peer Review (OPR) is, according to Ford (2015, p.2), 'any scholarly review mechanism providing disclosure of author and referee identities to one another at any point during the peer review or publication process'. OPR, available since at least the 90s (Tattersall, 2015), takes a variety of forms on a continuum of openness, pre- and post-publication, and with reviewers ranging from anyone who chooses to access and comment on a paper, who are journal subscribers, or reviewers allocated by the publication. Authors may be able to access and respond to reviews with author[s] and reviewer[s] named or degrees of openness may be optional with, for example, a reviewer able to opt for anonymity. Similarly reviews and author responses to those reviews may be published along with the publication.[22]

Criticising the 'blackbox' of traditional peer review for allowing reviewers, editors and authors[23] significant power to subvert the process, Ross-Hellauer (2017) offers a schema of seven OPR traits including 'open identities' (p.7), alternatively termed signed or unblinded review. While advocates for OPR point to its potential to improve accountability as well as to provide credit for the academic labour of reviewing, Ross-Hellauer (2017, p.9) also emphasises ways in which OPR might address some of our concerns about secret peer review, citing van Rooyen et al's (1999, p.23) claim that '… it seems unjust that authors should be "judged" by reviewers hiding behind anonymity'. However, studies conducted to date mainly in medicine, have shown no significant effects in favour of or against open identity reviewing and clearly further research is required before we could fully abandon secret identities in pursuit of reviews driven by the imperative 'to bring greater transparency, accountability, inclusivity and flexibility to the restricted traditional model of peer review' (Ross-Hellauer, 2017, p. 11).

That imperative is echoed in the first editorial of a new journal, *The Public Philosophy Journal* (PPJ[24]). Meeting some of our calls to recast peer review and to premise it on a more open, dialogic, and deliberative model, the PPJ uses 'Formative Peer Review, 'a structured form of peer engagement rooted in trust and a shared commitment to improving the work through candid and collegial feedback'[25]. This form of open identity peer review involves a deliberative open process with authors submitting drafts and nominating a reviewer in a review process explicitly intended to enhance the paper by encouraging responsiveness between the review participants with reviewers and reviewees working together, and certainly not in secret.

Yet some may find such OPR processes too radical a response to the concerns we have raised here. OPR is not yet widely accepted by the academic community although a recent survey report (2015[26]) suggests support is increasing. We have argued that defence of anonymised secret peer review on the grounds that it affords impartiality is flawed, but such anonymity is also often defended on the grounds that its secrecy protects reviewers from retribution by reviewees. Pointing to this as a prominent reason for resistance to more open forms of review, Flier (2016, p. 2) notes that this 'is hardly complimentary to the scientific community' and 'brings shame on a profession committed at its core to the pursuit of knowledge and truth'. This is a challenging

rebuttal to this oft cited defence of anonymised peer review. It coheres, however, with our argument that we should interrogate the assumptions underlying any and all forms of peer review and, with respect to secrecy, that the defences that so frequently buttress anonymised peer review should be questioned. Putting philosophical scrutiny to work to question the intentions and consequences of secrecy in peer review does not, however, entail that we dismiss the realities of today's competitive culture. It is perfectly feasible to conjecture that, pre- or post-publication, some of us have provided more favourable reviews if there was any likelihood that we could become known to the reviewee than those we might have given behind a cloak of secrecy. Hence it would be naive to entirely dismiss reviewer protection as a defence for anonymised review in an academy beset by performativity, audit and accountability. However, we do accept Flier's (2016) call that we should not regard reviewee protection 'as unavoidable' and to be countered by secrecy 'and its adverse consequences'. The wider problem should be addressed. That problem is an issue of power and often of patronage. It goes far deeper than peer review but against that background we offer a number of additional suggestions that might go some way to addressing the most pernicious elements of secrecy.

A practical recommendation that emerges from our exploration of peer review pertains to training as part of a shift towards more open peer review. Peer review training has increased significantly in the last few decades with a multiplicity of online resources available[27] and yet research, conducted mainly in medicine and the sciences, questions such training and, in some cases, points to no evidence of positive effects. This is summarised in a recent blog entry in which Khoo (2018) concludes that 'peer reviewer training is not going to save peer review or populate publisher databases with high-quality reviewers' but that it may enable novice reviewers to 'feel more confident or give them a chance to network within their discipline'. While Khoo alludes here to networking as a positive spin-off from training, we position this not as a spin-off but as an integral element of the peer review process, re-purposing networking as a strategy to offset the negative aspects of secrecy while enhancing peer review outcomes. The excellent technology-enabled peer review training already available might then become a component rather than an end of ongoing enhancement of peer review. That enhancement would be located in a collegial supportive environment in which deliberation with peers could address the challenges of secrecy with particular respect to bias and the dangers of inexpert judgments. Dialogue would be at the heart of such

developments, not only supporting novice reviewers but providing a space for even the most experienced reviewers to render explicit their thinking, their experience and their judgments. While not dissenting from Callaham and Tercier's (2007, p.32) insistence that the requisite skills of peer review 'may be as ill defined and hard to impart as is "common sense" ', we do not dismiss skills training. Rather we locate it as one element of a broader deliberative approach designed to ameliorate the worst tendencies of secret peer review.

Prompted by the urgency and proliferation of anonymised peer review in preparation for the REF we have argued that the so-called gold standard of peer review is tarnished. In focussing on dimensions of secrecy and its intentions and consequences we have sought to rupture what is, arguably, the unquestioning compliance that we, reviewers and the reviewed, exhibit at a time when peer review has far-reaching consequences for us all. As philosophers of education, we have an opportunity to engage with possibilities for forms of that review that do not merely accept that it is 'the best we have' but that strive to make it better. Such possibilities include the requirement to question the need for secrecy and to make peer review a more open dialogic activity premised on careful deliberation as part of both the review process and ongoing training for that process. Rather than continuing to accept the status quo, such possibilities could more aptly reflect our aspirations to sustain an academic community. Following the Philosophy of Education Society of Great Britain (PESGB[28]) this would promote 'the philosophy of education in a climate of inclusion, tolerance and respect for diversity'.[29] This entails bringing the challenges of peer review, especially its most secretive enactments, into the open, into debate with colleagues across the field of education.

While REF2021 renders scrutiny of peer review urgent it also provides, if we are willing to stop and question it, an opportunity for deliberation. That deliberation will not provide easy answers and we do not proffer a conclusive solution. Rather we have challenged the credentials of anonymous review, arguing that it could be made less secretive if we start from the premise that secrecy is no guarantee of impartiality or quality. The first onus is on those who insist on secrecy to more openly defend the rationale for that secrecy. The second onus is on all of us engaging in and judged by secretive peer review to question its intentions and to be aware of its far reaching consequences. Finally, and to return to our epigraph, the secrecy of peer review should

surely not be a riddle: 'a puzzle or joke in which you ask a question that seems to be nonsense but which has a clever or amusing answer'[30]. Anonymised peer review should not be a puzzle. Its effects are far too important for it to be a joke and as it continues to shape our academic identities, peer review requires open deliberation, both in and about the process.

REFERENCES

Allmer, T. (2018) Theorising and analysing academic labour, *tripleC: Communication, Capitalism & Critique*, 16(1), 49-77, http://www.allmer.uti.at/wp-content/uploads/2011/11/868-3905-2-PB1.pdf, last accessed 14/08/2018.

Baez, B. (2002) Confidentiality and peer review: the paradox of secrecy in academe, *The Review of Higher Education*, 25(2), 163-183.

Benhabib, S. (1996) Toward a deliberative model of democratic legitimacy, in S. Benhabib (ed) *Democracy and Difference,* Princeton: Princeton University Press, 67-94.

Biagioli, M. (2002) From book censorship to academic peer review, *Emergences: Journal for the Study of Media & Composite Cultures*, 12(1), 11-45.

Berry, R.L. (1980) Academic freedom and peer reviews of research proposals and papers, *American Journal of Agricultural Economics*, 62(4), 639-646.

Bohman, J. (1998) The coming of age of deliberative democracy, *Journal of Political Philosophy*, 6(4), 400-425.

Bok, S. (1989) *Secrets: On the Ethics of Concealment and Revelation.* New York: Pantheon Books.

Butchard, D., Rowberry, S., Squires, C. & Tasker, G. (2017) *Peer Review in Practice*, London: UCL Press, https://ucldigitalpress.co.uk/BOOC/Article/1/57/, last accessed 14/08/2018.

Button, M.E., & Garrett, J. (2016) Impartiality in political judgment: deliberative not philosophical, *Political Studies* 64(IS), 35-52.

Callaham, M.L. and Tercier, J. (2007) The relationship of previous training and experience of journal peer reviewers to subsequent review quality, *PLoS Med* 4(1): e40, https://doi.org/10.1371/journal.pmed.004004, last accessed 14/08/2018.

Conroy, J. & Smith, R. (2017) The ethics of research excellence, *Journal of Philosophy of Education*, 51(4), 693-708.

Derricourt, R. (2012) Peer review: Fetishes, fallacies, and perceptions, *Journal of Scholarly Publishing*, 43(2), 137-147.

Dryzek, J. (2001) Legitimacy and economy in deliberative democracy, *Political Theory*, 29 (5), 651-669.

Flier, J.S. (2016) It's time to overhaul the secretive peer review process, https://www.statnews.com/2016/12/05/peer-review-process-science/, last accessed 14/08/2018.

Ford, E. (2015) Open peer review at four STEM journals: an observational overview, *F1000Research,* 4 (6), https://f1000research.com/articles/4-6/v2, last accessed 14/08/2018.

Friedman, M. (1993) The impracticality of impartiality, in *What are Friends for? Feminist Perspectives on Personal Relationships and Moral Theory*, Ithaca: Cornell University Press, 9-34.

Gould, T.H.P. (2010) Scholar as e-publisher - the future role of [anonymous] peer review within online publishing, *Journal of Scholarly Publishing*, 41(4), 428-448.

Hicks, D. (2012) Performance-based university research funding systems. *Research Policy*, 41(2), pp. 251–261.

Habermas, J. (1979) Moral development and ego identity, in *Communication and the Evolution of Society* [trans. T McCarthy], Boston: Beacon Press.

Hirsch, J.A., Manchikanti, L., Albuquerque, F.C., Leslie-Mazwi, T.M., Lev, M.H., Linfante, I., Mocco, J., Rai, A.T., Schaefer, P.W. & Tarr, R.W. (2017) The peer review process: a primer for JNIS readers, *Journal of NeuroInterventional Surgery*, 9(e3-e6), https://jnis.bmj.com/content/9/e1/e3, last accessed 14/08/2018.

Jackson, L., Peters, M.A., Benade, L., Devine, N., Arndt, S., Forster,D., Gibbons, A., Grierson, E., Jandrić, P., Lazaroiu, G., Locke, K., Mihaila, R., Stewart, G., Tesar, M., Roberts, P. & Ozoliņš, J. (2018) (2018) Is peer review in academic publishing still working?, *Open Review of Educational Research*, 5(1), pp. 95-112, https://doi.org/10.1080/23265507.2018.1479139, last accessed 14/08/2018.

Kernohan, D. (2018) What we now know about REF2021, https://wonkhe.com/blogs/what-we-now-know-about-ref2021/, 23rd July, last accessed 14/08/2018.

Khoo, S. (2018) There is little evidence to suggest peer reviewer training programmes improve the quality of reviews, http://blogs.lse.ac.uk/impactofsocialsciences/2018/05/23/there-is-little-evidence-to-suggest-peer-reviewer-training-programmes-improve-the-quality-of-reviews/ last accessed 14/08/2018.

Lucas, L. (2017) Evaluating academic research: ambivalence, anxiety and audit in the risk university, in S. Wright and C. Shore (eds) *Death of the Public University?: Uncertain Futures for Higher Education in the Knowledge Economy*, Oxford: Berghahn Books, pp.213-218.

McCulloch, S. (2017) Hobson's choice: the effects of research evaluation on academics' writing practices in England, *Aslib Journal of Information Management*, 69 (5), 503-515.

Miller, B., Pevehouse, J., Rogowski, R., & Tingley, D. (2013) How to be a peer reviewer *PS: Political Science and Politics* 46 (1), 120-123.

Murphy, T. (2017) Revising the Research Excellence Framework: Ensuring quality in REF2021, or new challenges ahead?, *Perspectives: Policy and Practice in Higher Education*, 21(1), 34-39.

Power, M. (1997) *The Audit Society: Rituals of Verification*, Oxford: Oxford University Press.

Ross-Hellauer, T. (2017) What is open peer review? A systematic review, *F1000Research*, 6(588), https://f1000research.com/articles/6-588/v2, last accessed 14/08/2018.

Shatz, D. (2004) *Peer Review: A Critical Inquiry,* New York: Rowman & Littlefield.

Smith, R. (2006) Peer review: a flawed process at the heart of science and journals, *Journal of the Royal Society of Medicine,* 99(4): 178-182.

Tattersall, A. (2015) For what it's worth - the open peer review landscape*, Online Information Review*, 39(5), 649-663.

UK Government (2014) *Building on success and learning from experience - an independent review of the Research Excellence Framework* (referred to as the Stern Review), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/541338/ind-16-9-ref-stern-review.pdf. last accessed 14/08/2018.

Van Noorden, R. (2013) Open access: The true cost of science publishing, *Nature*, 495(7442), 426-9.

Van Rooyen, S., Godlee, F., Evans, S, Black, N. and Smith, R. (1999) Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial, *British Medical Journal*, 2(318), 23-27.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J. and Johnson, B. (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*, http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/Independentresearch/2015/The,Metric,Tide/2015_metric_tide.pdf, last accessed 14/08/2018.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L. de Rijcke, S., Rushforth, A. & Franssen, T. (2015) *The Metric Tide: Literature Review* (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management), HEFCE, http://www.dcscience.net/2015_metrictideS1.pdf, last accessed 14/08/2018.

Young, I.M. (1990) The ideal of impartiality and the civic public, in I.M. Young, *Justice and the Politics of Difference*, Princeton: Princeton University Press, 96-121.

---

[1] We do not focus on peer review for grant applications but our argument is likely relevant to anonymous review of such applications.

[2] The UK REF is 'a process of expert review, carried out by expert panels' with 'three distinct elements … the quality of outputs (e.g. publications, performances, and exhibitions), their impact beyond academia, and the environment that supports research' from http://www.ref.ac.uk/about/whatref/. We focus here on outputs.

[3] See http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101043/kj1a27837enn.pdf

[4] We follow Allmer (2017, p.56) in using academic labour rather than academic work as 'universities are part of capitalism and academics are embedded into class relations'.

[5] 'Outputs' form 60% of the overall submission to any Unit of Assessment.

[6] See, for example, the British Educational Research Journal, the most frequent outlet overall for Education in the 2014 REF, https://onlinelibrary.wiley.com/page/journal/14693518/homepage/forauthors.html.

[7] Single blind review, by contrast, entails that editors and referees know the identities of authors while the identities of reviewers remain concealed to authors.

[8] Some journals make other reviews available to reviewers once submitted and, often, when a publication decision has been made.

[9] See https://orgtheory.wordpress.com/2007/01/23/a-modest-proposal-triple-blind-review/ for a proposal for, at least, triple blind review in which the editor is not aware of the author's identity.

[10] In Scotland, the funding stream related to REF is the Research Excellence Grant (REG), in England Quality-related Research funding (QR). This funding was, following REF2014, allocated only to 3* or 4* outputs.

[11] On average, this means 2.5 outputs per eligible staff member.

[12] http://www.ref.ac.uk/publications/2018/draftguidanceonsubmissions201801.html

[13] Final decisions on REF2021 will not be available until Winter 2018-19.

[14] See, for example, Van Noorden (2013) on 'free' peer review labour provided to publishers who then charge significant sums for access to their journals.

[15] See Murphy (2017) for gaming still possible in REF2021 and, on REF2014, Marginson's (2014) blog entry on 'Research excellence: getting better all the time – or is it?', at https://ioelondonblog.wordpress.com/2014/12/19/research-excellence-getting-better-all-the-time-or-is-it/

[16] "The Worst Piece of Peer Review I've Ever Received," *Times Higher Education*, August 6, 2015, https://www.timeshighereducation.com/features/the-worst-piece-of-peer-review-ive-ever-received.

[17] See https://www.sciencealert.com/these-8-papers-were-rejected-before-going-on-to-win-the-nobel-prize

[18] https://www.elsevier.com/en-gb/reviewers/what-is-peer-review

[19] See https://www.hepi.ac.uk/2018/01/04/guide-uk-league-tables-higher-education/

[20] As with our experience submitting a paper to a well-known journal.

[21] See, for example, https://www.ecu.ac.uk/wp-content/uploads/2014/07/unconscious-bias-and-higher-education.pdf , https://royalsociety.org/topics-policy/publications/2015/unconscious-bias/, https://implicit.harvard.edu/implicit/, https://www.tolerance.org/professional-development/test-yourself-for-hidden-bias.

[22] See Ford (2015) for an example of OPR with reviewers/reviewed identities disclosed and reviews/responses published.

[23] For example, authors can subvert the secrecy of their identities by, for example, self-citation.

[24] http://publicphilosophyjournal.org/

[25] http://publicphilosophyjournal.org/about/review/ and see http://publicphilosophyjournal.org/about/review/ for details of PPJ peer review processes.

[26] Mark Ware Consulting (2016) Publishing Research Consortium Peer Review Survey 2015, http://publishingresearchconsortium.com/ with 50% of respondents accepting open review but an ongoing reluctance to espouse signed reviews published with papers.

[27] See, for example, http://peerreviewtraining.esrc.ac.uk/ and https://authorservices.wiley.com/Reviewers/journal-reviewers/becoming-a-reviewer.html/peer-review-training.html, and https://hub.wiley.com/community/exchanges/discover/blog/2014/11/07/mentoring-the-next-generation-of-reviewers

[28] The parent society of this journal.

[29] http://www.philosophy-of-education.org/

[30] https://www.collinsdictionary.com/dictionary/english/riddle