PERSPECTIVES

Peer review
Objective screening or wishful thinking?

Ken Hyland
University of East Anglia

1. Introduction

Despite the massive changes in academic publishing in recent years, one thing remains more or less constant: the disdain many academics feel for peer review, and perhaps for peer reviewers themselves. This is possibly the most contentious and secretive practice in our academic lives, reviled and tolerated in equal measure but rarely loved. Slow, biased, contradictory, hurtful or wilfully obtuse, reviewers come in for a lot of stick. But are the criticisms levelled at peer review merited or is it an effective means of separating the sweet from the sour? Here I argue that while there is a lot wrong with peer review, it serves an important function for the academic community and that ERPP teachers have a role to play in making it work.

2. What’s it for?

Prestigious journals stake their reputations on the quality of the papers they publish and therefore on the standards of their peer review, leading to rejection rates in the top humanities journals of over 90% (APA, 2018). While rejections create deep passions among academics, peer review contributes to the creation and archiving of knowledge in several ways.

For readers, it acts as a filter. With more than a million papers published each year, academics can only keep abreast of significant work if trusted reviewers pre-read the literature. For authors, reviewers’ reports provide an indication of how another researcher understands their work and its contribution. It is the guidance of useful feedback which is usually responsible for bringing ‘Revise and Resubmit’ papers to publication. In fact, because most papers eventually find a home in a journal somewhere, peer review may be regarded as a mechanism for deciding where a paper is published rather than whether it is published.
Most centrally, however, peer review helps screen submissions for publication, providing the expertise and time that editors themselves might lack. Unreviewed research posted on personal websites is regarded with scepticism by the academic community. In their study of 4000 researchers, for example, Mulligan, Hall, and Raphael (2013) found that most respondents saw peer review as “the most effective mechanism for ensuring the reliability, integrity, and consistency of the scholarly literature” (p. 149).

In fact, peer review underpins how academia sees itself and, indeed, embodies two of Merton’s (1973) four norms defining scientific practice. Peer review contributes to Universalism, or an adherence to objectivity rather than personal self-interest, and to Organised scepticism, so that no theory is accepted merely on the authority of the proponent (Howard, 2012). It is often seen as the cornerstone of academic credibility and fundamental to the development and integration of new research (Hyland, 2015). More generally, of course, it reaches deep into the lives of individual researchers so that appearing in peer-reviewed publications has become the benchmark for ranking scholars in a very high stakes game. Publication recognises a paper’s worth and an academic’s credibility.

3. What’s not to like?

Well, plenty it seems, so, for example, the authors of a recent paper believe that peer review is plagued by:

- a preponderance of incompetent reviewers,
- a lack of constructive criticism and the maintenance of orthodoxy,
- relative ease in identifying blinded authors,
- editorial passivity,
- and long waits to receive reviews. (Sciullo & Duncan, 2019, p. 248)

I’ll discuss the main issues under five areas: consequences, delays, bias, disrespect and skulduggery.

3.1 Reviewer decisions (really) matter

While no one enjoys having their work critically dismembered, it can be the impact of reviewer comments which are most devastating. These “mysterious and intimidating figures” (Tardy, 2019), masked by anonymity, have the power to influence our professional lives as careers can hang on their decisions. We live in an accountability culture that evaluates individuals and their research by where papers are published and how frequently they are cited. Peer review is the most visible manifestation of this and can influence appointments, promotions, and funding.
Publishing now drives research, rather than the other way around, as academics across the globe must now publish, usually in English, to present their ideas, increase their visibility and navigate their careers. Consequently, there are as many as eight million scholars working in 17,000 universities worldwide seeking to publish in English-language journals each year (Reller, 2016). The number of papers is also staggering. Academics now produce over 2.5 million new peer reviewed articles a year (Ware & Mabe, 2015) with the global scientific output doubling every nine years (Bornmann & Mutz, 2014). So the pressure to publish has never been greater, contributing to levels of psychological distress among academics higher than those of front-line police and staff in hospital accident units (Grove, 2018).

3.2 Long decision delays

A common complaint levelled at peer review journals is the length of time it takes to get a decision, often taking far longer than the 3 months journals usually promise. The main reason for this is not (only) the tardiness of editors or reviewers, but the difficulties of finding reviewers. I once, as co-editor of Applied Linguistics, sent a submission to 13 people before I found three who would do it. All journals encounter increasing difficulty in identifying willing readers and this refereeing burden, moreover, falls particularly heavily on senior academics and disproportionately on US researchers (Warne, 2016).

The scale of the publishing industry puts a huge strain on the creaking peer review system. The journal Nature alone receives over 10,000 manuscripts a year, for instance, and Elsevier’s online submission system processes more than 120,000 new manuscripts per month. This explosion of submissions, of course, is largely because of the reward system. As the Chronicle of Higher Education has observed:

Physica A publishes 3,000 pages each year. Why? Senior physics professors have well-financed labs with five to 10 Ph.D.-student researchers. Since the latter increasingly need more publications to compete for academic jobs, the number of published pages keeps climbing.

(Bauerlein, Gad-el-Hak, Grody, McKelvey, & Trimble, 2010)

This surge in submissions demands more reviewers. So, just one publisher, Elsevier, made use of 700,000 peer reviewers in 2015 alone to conduct 1.8 million reviews (Reller, 2016). But while the population of researchers is growing, the pool of potential reviewers may be shrinking, certainly the proportion of writers to reviewers is shifting significantly, so that Chinese authors, who are among the most prolific of submitters for example, publish twice as much as they review.
This requirement for ever-increasing numbers of reviewers has also occurred alongside an ever-narrowing coverage of journals. Specialism generates more new journals so, for instance, there are now 61,300 dentistry journals listed in Pubmed (National Library of Medicine, 2019) and over 700 in language and linguistics (SJR, 2017). This increases the demand for reviewers with specific expertise, making it harder for editors to find suitable readers.

Reviewer shortages are, in addition, happening at a time when reviewing has become a marginalised part of an academic’s role, so that even willing reviewers are often forced to do a hurried job. Reviewing is done without reward or credit, and so time for it suffers as universities demand more teaching, more admin, more outreach, more research and more everything else. Given these competing pressures, many academics feel that the compensation offered for reviewing, such as journal acknowledgments, positions on editorial boards, free journal access, etc. is just not worth it. Editors therefore turn to untried individuals and the quality of reviews may suffer.

3.3 Bias and subjectivity

Peer review bias is a violation of impartiality in the evaluation of a submission (Haffar, Bazerbachi, & Murad, 2019) and so is difficult to completely avoid. Reviewers are academics themselves and have pet theories and an investment in certain approaches and ideas. Research has therefore turned up evidence of ‘confirmatory bias’, for example, showing that reviewers are less likely to criticize work consistent with their own views (Ernst & Resch, 1994). Novel ideas and approaches thus bear a greater burden of proof than those following existing paradigms. Peer review bias is difficult to detect, however, and reviewers may even introduce bias inadvertently by asking authors to modify analyses, perform post hoc explanations, and so on.

Authors also see bias in conflicting reviews. These are so common that reviewer agreement is little better than chance (Rothwell & Martyn, 2000) and it is just not feasible to get the six reviews for every paper needed for statistical reliability (Fletcher & Fletcher, 2003). Reviewers can only bring their knowledge of what the community believes and might accept by seeing how far the paper establishes novelty and aligns this with an existing literature. The trouble is that the literature grows so rapidly that we can never know when a reviewer might see what we have disregarded as poor scholarship. Editors, however, often value these different reviewer perspectives and see a review as an argument for a particular judgement.

Many EAL authors also feel that they are unfairly penalised by reviewers for non-standard language use (e.g. Perez-Llantada, 2014). It is true that reviewers
frequently comment on poor grammar, style or rhetorical conventions, but there is little evidence of systematic bias against NNESs. Research suggests that reviewers do not typically consider the non-native speaker status of authors in making decisions and the quality of the language is rarely a decisive factor in rejection (Belcher, 2007; Coniam, 2012). More important is the argument and the need to present the originality and value of the research. There may even be greater tolerance for non-standard language as growing numbers of non-native speakers submit papers and take on reviewer and editing roles themselves (Hyland, 2016). Rozycki and Johnson (2013), for example, found widespread use of simplified grammar in prize-winning papers by non-native English speakers in in IEEE journals, engineering’s most prestigious publications.

It is true, however, that there may be a preference for articles from the US and Canada (Okike, Kocher, Mehlman, Heckman, & Bhandari, 2008; Ross, Gross, Desai, Hong, Grant, & Daniels, 2006) while good submissions from countries with low Gross Domestic Product are less likely to be accepted (Saposnik, Ovbiagele, Raptis, Fisher & Johnston (2014). Double-blind review helps eliminate much of this (Ross et al., 2006), but the high degree of specialization in modern science makes anonymity difficult. The small size of some fields, highly specialised journals, pre-publication presentations or blogs, and self-citation mean reviewers can sometimes guess the author’s identity.

We have to remember that reviewers are not computers running algorithms which check manuscripts against objective criteria. Values and beliefs always feed into recommendations. Far from being surprising, then, disagreements and subjectivity are inevitable: an inescapable aspect of scientific fact construction.

3.4 Condescension and rudeness

Reviewer comments can improve a submission but not all reviewers are mentors, working with authors to bring manuscripts to publication. Some reviews are caustic and unhelpful, as examples on the *shit my reviewers say* website\(^1\) testify:

This work is antithetical to the spirit of research and will impede potentially important developments.

I am personally offended that the authors believed that this study had a reasonable chance of being accepted to a serious scientific journal. When the reader is finished struggling through all the methods and results, he/she is left wondering whether it was worth the time.

---

Clearly reviewers’ reports can be demoralizing and while anonymity might help prevent personal bias, it can make reviewers less accountable (e.g. Waggenknecht, 2018). Certainly, analysis of reviews show considerable use of attitude markers and self-mention to emphasise the reviewer’s certainty and authority (Paltridge, 2017). This can be extremely demotivating, as Kwan (2013, p.213) observes:

Many first-time writers are confused, discouraged or even shocked by the negative reviews they receive and the substantial revisions requested...Some never attempt to revise and resubmit their work that reviewers see as having potential for publication.

But not only novices get rejections, Cassey and Blackburn (2003) found that about one fifth of submissions by top researchers in ecology are rejected the first time around. Nor is this the end. Calcagno et al.’s (2012) study of over 80,000 rejected bioscience papers found that these rejected papers received significantly more citations when they eventually got accepted elsewhere than those that were accepted on first submission. Criticism is, of course, a key feature of reviews, comprising perhaps half of all comments (Fortanet, 2008). Bluntness may be due to a hurried report, personal style or even a lack of pragmatic competence, as EAL reviewers seem far more direct than native English speakers (Paltridge, 2017). In my corpus of 150 peer reviews from leading applied linguistics journals, however, directness is rare, and both criticisms and suggestions are largely hedged (Hyland, 2015). A good news/bad news macrostructure is widely used while praise tends to be global and criticism more specific (see also Belcher, 2007, p.9). Generally, reviewers seek to create a positive, sympathetic relationship with writers.

As a journal editor for over 16 years, I occasionally edited or binned unreasonably harsh reviews, but the majority of reviewers seem to take considerable care to frame negative reviews in a helpful, collegial way. We have to remember that reviewers also submit research to journals and can empathise with authors, imagining the kinds of comments they would like to receive themselves. This encourages a certain collegiality and a tendency to be as supportive as possible to encourage authors to keep working to get published.

3.5 Skulduggery and dishonesty

Critics of peer review further worry that reviewers may favor their friends or themselves, that they may review their competitors’ work unfairly or that they will take advantage of ideas they find in the papers they review. Science is a competitive world and the rewards of success mean misconduct occurs, but such problems seem to be rare (e.g. Godlee, Gale, Martyn, 1998; Ward & Donnelly, 1998).
While peer review is probably not a nest of dishonest practice, it is very poor at stopping dishonesty in others. A majority of the 4000 reviewers questioned by the Sense about Science (2009) survey thought that peer review should detect plagiarism (81%) or fraud (79%) but fewer (38% & 33%) think it is capable of this. Reviewers can generally comment on the correspondence between claims and data, point out oversights, suggest alternative explanations, highlight design flaws, etc, but are hopeless at spotting fabricated data or plagiarism (Brainard & You, 2018). The ever-growing number of papers and reviewer workloads make it difficult to keep up with the field and find time for careful reviewing so questionable work can find its way to publication. Retractions of articles remain relatively rare, however, with only about four in every 10,000 papers affected. Ironically, about 20% of articles which are pulled from journals involve fake peer reviews (Huag, 2015), thus reinforcing respect for the practice!

4. So what’s to be done?

We need more, better trained and better rewarded reviewers to cut waiting times, reduce bias and rudeness and generally make the whole thing more accountable. Beyond, that, we need to overhaul a dysfunctional system which puts unreasonable pressure on authors, reviewers and editors in a high stakes numbers game while generating huge profits for publishers. Attempting to quantify knowledge, financially reward publication and judge individuals by the length of their bibliographies is neither ethical nor sustainable. There are, however, more immediate things that can be done and ERPP teachers can help.

4.1 Train peer reviewers

While academic writing is taught to undergraduates, reviewing is rarely covered in PhD programmes or ERPP courses. In a questionnaire study of 45 reviewers, for example, Paltridge (2013) found over half had learnt to do reviews by reading reviews of their own submissions and the rest had learnt to write by just doing them. This might help explain the results of a recent study of editors and reviewers in software engineering that two thirds of reviews were considered misleading or merely adequate (Prechelt, Graziotin, & Méndez Fernández, 2017).

Peer review training might not only help improve the quality of published research, it could also furnish early career scholars with analytical and communication skills to contribute more effectively to their disciplines. In fact, the UK House of Commons Science and Technology Committee (2011) recommended that researchers be offered such peer review training. In the medical sciences, *The
British Medical Journal and Annals of Emergency Medicine offer training to their reviewers through workshops and materials on their websites which walk reviewers through evaluating a manuscript and give examples of good and bad reviews. In contrast, BioMed Central has a peer review blog, covering issues like how to decide whether to take on a particular review, how to approach it, and how to re-review.

Several experimental studies suggest training packages have little impact on the quality of reviews (Callaham & Tercier, 2007; Schroter, Black, Evans, Godlee, Osorio & Smith, 2008), but it is difficult to transfer these results to real-world practice (Jefferson, Rudin, Brodney Folse, & Davidoff, 2007). A survey of 3000 reviewers found that most reviewer training comes in the form of journal guidelines and collegial advice, but the vast majority wanted formal training (Warne, 2016). There is certainly opportunities for supervisors or faculty advisors to assist in this regard, while giving feedback to writers is the ERPP teacher’s stock in trade. Surely there is a role for EAP and ERPP instructors here, analysing reviews, researching practice and helping novices, and journals, to unambiguously formulate expectations of ‘good reviews’. This might also have the effect of mitigating harsh reviews and encouraging reviewers to be constructive.

Whatever the source of the training, it seems crucial to have some experiential element as a component so that reviewers can see the potential pragmatic impact of comments on authors. This would minimise the kind of feedback which Becher and Trowler (2001, p.89) describe as “venting of personal preferences or antipathies” or which, more dramatically, Suls & Martin (2009, p.302) calls “bloodletting”. Paltridge (2017, p.154) argues that training needs to be hands-on and reflective, under the guidance of an experienced reviewer. They might require, for instance, trainees to read examples of reviewers’ reports and critique them for what they find most and least useful in the reviewers’ comments. They can then create their own list of issues to focus on in the writing of reviews based on the observations they make from reading the reports.

It is also important that authors come to understand the process of peer review, not only as a key element of getting their work published, but to involve them in the process and avoid the false author/reviewer dichotomy that seems to be common among some junior scholars.

4.2 Reward peer reviewers

Reviewers often feel put upon by editors as the work is unpaid, unrewarded, and often criticized by authors. Would the pool of reviewers increase if they were rewarded for it?

Two large scale surveys of over 6000 academics in each, show that reviewers typically spend five hours on each review and overwhelmingly do this because they want to contribute to the academic community and improve new papers (Publishing Research Consortium, 2016; Taylor & Francis, 2015). Others, particularly early career researchers, may feel motivated by being trusted to comment on a peer’s work by an editor, to be a part of professional debates and to gain experience which may improve their own writing. Other researchers see the time spent reviewing as more efficient than having to comb through non-reviewed papers and by getting access to a literature filtered by the submitting author.

Publishers have long sought to reward reviewers in small ways such as annual ‘thank you’ lists naming reviewers on journal websites or offering free online access to the journal for 6 months. Beyond this, Elsevier has offered reviewers certificates for their work and Nature trialled the idea of putting reviewers’ names on published papers to acknowledge their contribution. PLOS, the respected Open Access science publisher, also allows authors to publish their peer review history with reviewers’ names attached. Reviewers have also recently been permitted to opt-in to have their peer review activity posted on their ORCID account, and because the record does not contain details about the manuscript this allows credit to be gained while retaining anonymity. Another idea, promoted by Publons, allows researchers to collect digital badges which reviewers might use for public and institutional recognition for their work, enabling junior researchers to submit them as part of a tenure portfolio, for example. So far, publishers have been slow to discuss financial appreciation.

Reviewers strongly believe, however, that reviewing is inadequately acknowledged and should carry more weight in institutional evaluation processes (Warne, 2016). Publishers save over £1.9 billion a year in free reviews (Research Information Network, 2008) and they should start transferring some of this to reviewers. Some small Open Access Publishers such as Veruscript in the UK have begun offering reviewers modest payments based on the fees they charge authors, but the major players have been strangely quiet on the issue. There is no guarantee, however, that payment will improve the quality of peer review and it might change how the task is seen, replacing more altruistic reasons for doing the work. The two surveys discussed earlier suggest that respondents would strongly prefer institutional recognition in performance appraisals to hard cash. So, if institutions make it clear that peer review is a pathway to progression, tenure and funding,
researchers will prioritize it. Similarly, if world ranking bodies weighted peer review contributions, then universities would have greater incentive to reward individuals for it (Preston & Culley, 2017).

Benchmark standards are obviously necessary, but the technology exists to record and track contributions of an individual’s review record. Platforms like Publons and ScienceOpen already do this by linking the reviewing records of individuals for a particular journal into ORCID. Performance metrics connected to career recognition can thus encourage and certify peer review.

4.3 Restructure peer reviewing

Various options have been proposed to replace current practices. To improve accountability, some critics advocate an open peer review system, where both authors and reviewers know the identity of each other. This might increase transparency, reduce abusive comments, encourage reviewers to be constructive, and possibly prevent plagiarism. (O’Connor, Cousar, Lentini, Castillo, Halm, & Zeffiro, 2017). The downside is the possible creation of animosities and retaliation between authors and reviewers (DeCoursey, 2006). The British Medical Journal (BMJ) adopted this system some 20 years ago, and all PLOS journals recently began offering authors the option to publish their peer review history alongside their accepted manuscript. This package includes the editor’s decision letter, reviewer comments and authors’ responses for each revision and allows reviewers to earn citations if they sign their names to their reviews. Nature and PLOS Medicine, however, have had problems with open peer review because of higher refusal rates, increased delays and non-engagement of authors and reviewers (Lee, Sugimoto, Zhang, & Cronin, 2013).

More radical is the idea of post-publication peer review (PPPR) which expands comments beyond a few experts to the judgement of a broader audience. In this model papers are published immediately after a light check by an editor then appointed referees and anyone else contribute reviews which are published next to the article with the name and affiliation of the reviewer. Authors can then either write rebuttals or try to address problems and reviewers can also address these problems. The interactive discussion deters authors from submitting low-quality manuscripts and allows reviewers to claim credit for their contribution, increasing motivation and perhaps encouraging better guidance in reviews (da Silva & Dobranszki, 2015).

This potentially progresses the current model by generating discussion, improving drafts and sharing information rapidly. Engaging outsiders in this way can strengthen research, but it runs the risk of operating like votes on YouTube and the possibility that only hot topics are taken up positively and dissenting
opinions are voted down. Lamont (2009, p.2) points out that “academia is not democratic,” as not everyone can properly evaluate originality and significance. Comments may become short and episodic, rather than comprehensive and conceptual, and dominated by the most aggressive. Nor is it clear whether already overloaded academics will take the time to comment on an even larger volume of published research, so that very few papers may attract comments.

The use of distributed ledger technology (DLT), which underpins bitcoin, could help make this work by taking reviewing away from individual journals and handing it to a centralised platform. This uses a universal database that is shared and synchronized across multiple sites without a central administrator or data storage. With open access platforms, DLTs can support PPPR as they make it easy to trace the origins and development of an article and for reviewers to be identified and compensated. This might make the process more transparent and encourage reviewers to be more constructive.

5. Some conclusions

The academic community has always argued that responsibility for the standard of published papers is a collective and broadly democratic one. It doesn’t depend on publishers, administrators or a senior elite of academics. There are serious strains on the system, however, and problems such as voluntary review, a shrinking pool of reviewers, mushrooming submissions, and unstandardized review criteria are all serious weaknesses. There are, then, many good reasons why the peer review process is broken. Peer review has become an easy target for frustrated writers, but as a journal editor who has given up a fair amount of my spare time to ask others to give up theirs to review new submissions, I find reviewer bashing slightly depressing.

We have to keep in mind that the purpose of peer review is not to ensure truth. It is a human process and subject to human imperfections. Equally, readers will draw their own conclusions about the value of the research and interpretations might be modified later by new data or paradigms. Simply, peer review helps ensure that the best possible version of research is published, and the practice needs to be judged on this.

A recent study shows that the more revisions a paper undergoes, the greater its citation impact (Rigby, Cox, & Julian, 2018), indicating that reviewers actively contribute to the text. Moreover, scholars seem committed to it. A massive 84% think that without peer review there would be no control in scientific communication and virtually all believe that their own papers were enhanced by peer-review (Publishing Research Consortium, 2016; Sense about Science, 2009).
most important question is how to improve it. By making review mandatory and rewarded and by separating it from journals many issues could be addressed. There is also a role for ERPP teachers, contributing our knowledge of academic texts and pragmatic interaction to peer review training to help ensure that the process is a formative, rather than a gatekeeping exercise and that it is conducted in a collegial and helpful way.

In the end, peer review is the value added to a paper by careful checking, but it is also the source of the academic community’s sense of fairness and self-reflection. Negotiating a paper to publication is a way, albeit often a fraught and distressing way, for writers to stay in the professional game.

References


Address for correspondence

Ken Hyland
k.hyland@uea.ac.uk