FEMS Microbiology Letters, 363, 2016, fnw115

doi: 10.1093/femsle/fnw115
Advance Access Publication Date: 25 April 2016
Commentary

COMMENTARY – Professional Development

Peer review: from recognition to improved practices

Pedro Cintas*

Department of Organic and Inorganic Chemistry, Faculty of Sciences—UEX, Avda de Elvas s/n, E-06006 Badajoz, Spain

*Corresponding author: Department of Organic and Inorganic Chemistry, Faculty of Sciences–UEX, Avenida de Elvas s/n, E-06006 Badajoz, Spain. Tel: +34-924-289-300; E-mail: pecintas@unex.es

One sentence summary: Peer review is still the most important pre-publication step to ensure high standards in scientific publishing, but further improvements along with public recognition should be developed.

Editor: Beatrix Fahnert

ABSTRACT

Scientific publishing has experienced profound changes in recent years, such as the advent of open-access journals, the increasing use of preprint archives or post-publication blogs, to name a few. One pillar still remains: peer review as a key ingredient that, in most cases, contributes to clarity and quality, often detecting errors and misinterpretations. Unfortunately, peer review is poorly recognized and good reviewers are rather a 'rare avis'. Even worse, this necessary task in science is generally overlooked in curricula and post-graduate education. Some considerations should help us all to ameliorate greatly our understanding and duties.

Keywords: peer review; scientific integrity; improved practices; peer review indexes; post-publication review; misconduct issues

INTRODUCTION

A few years ago I came across a diverse collection of quotes, ranging from hilarious to annoying, from reviewers on manuscripts submitted to *Environmental Microbiology* (Anon. 2010) and *Microbial Biotechnology* (Anon. 2011). In general, however, comments from reviewers are constructive enough helping authors to improve their contribution. Such comments also advise editors whether a paper in question should be accepted or rejected. This protocol lies thus in the heart of our currently accepted scientific practices. Peer review is not perfect, though it is surely the best tool to ensure clear, meaningful and error-free top research papers. As a matter of fact, however, peer review will hardly be able to unveil fraudulent practices, but it could also be unfair to blame for it, even if a paper becomes ultimately retracted (*vide infra*).

By definition, peer review involves the critical assessment of manuscripts submitted to journals (or published as individual chapters or monographs) by experts who are not part of the editorial staff (Hames 2007). The process is not aimed at perfection; rather it explores reliability without detracting from innovation and quality.

As strange as it may be, peer review is not new. The practice of annotating manuscripts by different workers or copyists dates back to ancient civilizations (Spier 2002). However, peer review as we know it can be attributed to Henry Oldenburg, the founder of the first scientific journal in the 17th century, *Philosophical Transactions of the Royal Society*. The practice was routinely adopted in the late 18th century by the Royal Societies of Edinburgh and London, which sought the advice of their members to select articles for publication. The system was gradually incorporated into different societies in Europe and America, but the procedures varied from journal to journal and from country to country. Peer review became standardized from the mid-20th century onwards (Burnham 1990; Kronick 1990).

The review process is therefore an indispensable element of scientific publishing. Without this filter and the clear-cut improvements the revision adds, it would certainly be impossible to maintain high standards. The review of manuscripts is much more crucial at present given the enormous flood of

Received: 5 March 2016; **Accepted:** 23 April 2016

papers, proceedings, preprints and news, mostly available through the internet, together with the increasing release of new journals, mostly adopting the open-access format. No one will be able to be updated enough, even in his/her narrow subject of interest. Under these premises, the bottleneck lies in the difficulty of finding qualified reviewers, especially in high-profile journals, capable of ascertaining the importance of the paper as a whole, paying attention to details, and providing a critical, yet fair and well-balanced assessment.

The first golden rule of peer review is that journal editors should ensure that both quality and ethical practices are met (Hames 2007). Getting ahead in editorship should not have to mean falling behind in authors' responsibility. It should be unnecessary to underline that every manuscript represents a collective task involving authors, editors, reviewers and funding agencies. However, editors of high-profile publications and those of open journals (these acquiring ever-increasing importance) often seek publicity and media coverage in addition to cutting-edge research (Silver 2014). One could wonder whether, after all, peer review exerts actually a discriminating role that drives to perfection. A conspicuous example is provided by Einstein's papers that appeared in 1905 in *Annalen der Physik*, which were never peer-reviewed. A better biosciences example is the Watson and Crick paper on the DNA structure, which was not sent for peer review. This decision might have been flawed because that model was not truly self-evident and even Watson and Crick had serious doubts about the correctness of their structure (Scher 2004). As noted by some experts in a comprehensive report on scientific publishing, commissioned by the British House of Commons (House of Commons 2011), peer review might ignore the significance of a seminal idea appreciated by contemporary scientists. This aspect has been recently illustrated by papers whose importance was not recognized for many years after publication, and have now proven to be extremely useful for emerging concepts or new applications (Ke *et al*. 2015).

In stark contrast, the downside is that big frauds overcome the vigilance of editors and reviewers, and are eventually unraveled by subsequent reinvestigation, often performed by anonymous researchers. Notable cases are that of Hwang Woo-suk, who reported to have succeeded in creating human embryonic stem cells by cloning, and that of Jan Hendrik Schön whose findings, if true, would have revolutionized the fields of materials and nanotechnology. The latter has been documented in detail in a monograph (Reich 2009); a detective work that highlights how fraud emerges and the roles of multiple protagonists, journals in particular, favoring a silent circle.

Overall, whether or not effective, peer review seems to be a necessary evil: its pluses largely exceed minuses. As mentioned above, fraudulent papers, i.e. containing false or fabricated data, often escape from peer review control. Moreover, such papers may be treated in a similar way to those reporting erroneous data or misinterpretations, which however, usually arise from honest research groups. Consider for instance, a paper reporting flawed results, which could be highlighted by a thoughtful reviewer. But what is rejected by one well-reputed journal can be found suitable for publication in subsequent second, third or fourth submissions going to second- or third-row journals with less strict refereeing. In a study appeared in 2012, Fang, Steen and Casadevall (2012) concluded that fraud had increased by *ca*. 10-fold since 1975, a higher trend than publication rates. Misconduct, and not errors, accounted for the majority of retractions (>67%), mostly in biosciences, which include cases of fraud, duplication and plagiarism. It is also true that high-profile journals have more retractions than low-impact journals, which simply tells us that cutting-edge research is more scrutinized than routine results. If results are important or far-reaching, there is no doubt that other laboratories will try their reproducibility. In striking constant, a large amount of papers across the globe dealing with variations on well-known facts remain substantially ignored.

MUCH ADO ABOUT IMPACTS

Numerous scientists have become obsessed by bibliometric indexes, such as the journal impact factors (IF) or h-indexes. The IF exerts a tremendous influence on the conduct of scientists, because funding agencies often use this criterion in decision making; i.e. IFs trigger a rewarding effect. As a prototypical example, the Research Excellence Framework (REF), a system for evaluating the quality of research in UK higher education institutions (http://www.ref.ac.uk), focuses on IFs as a metric to assess and score publications. Such agencies, and presumably most authors, ignore that the annual journal IF is largely due to a few selected and important papers, i.e. those receiving more citations. In a lucid analysis, microbiologists Casadevall and Fang (2015) recall that impact is not importance. They suggest some avenues to overcome this misconception such as to reduce the reliance on citation metrics for personal promotions and employment, or discuss the misuse of IFs in ethical courses. Unfortunately, the 'publish or perish' policy pervades contemporary science and it is really difficult to debunk this fallacy, especially among young scientists pursuing a career in academia or research.

Indexes are not harmful; only abuse and misuse make them ineffective. As once stated by Lord Kelvin, speaking of measurement, 'to measure is to know; if you cannot measure it, you cannot improve it'. Article metrics need to be analyzed to identify merit of each article and its authors. Metrics should move away from the number of papers published in journals with high IF to the true impact of one's research (House of Commons 2011).

Some indexes may be a necessity in any case, at least in terms of societal analysis and the responsibilities both editors and contributors (who are simultaneously authors and reviewers) acquire. In the context of retracted science, the introduction of a retraction index gives rise to further reflections (Fang, Casadevall and Morrison 2011). The retraction index for a given journal in a time interval (from 2001 to 2010 in that study) is the number of retractions multiplied by 1000 and divided by the number of published articles. This analysis shows a strong correlation between the journal retraction index and its IF. Once again, this trend reveals the lure and hype of IFs, rooted in our collective mind.

PEER REVIEW INDEXES

In line with the preceding analysis, the introduction of a peer review index turns out to be an inevitable metric. A so-called referee factor (Wilson 2006; Wilson and Lancaster 2006), defined as the sum of the IFs for the respective journals multiplied by the number of articles reviewed, has been proposed as an incentive for conducting this task and for assessments of professional performance. This referee factor shows some pros and cons. One editor, for instance, could send manuscripts to close colleagues and serving himself/herself as referee so long as this activity counts. Likewise, young scientists may be disfavored relative to senior authors in obtaining a referee factor (Rousseau 2006).

Moreover, further corrections should be introduced to avoid simplistic treatments. More than one review is usually required before manuscript acceptance and referees evaluate papers of different length and difficulty, so that the reviewing effort will be greater than the number of peer-reviewed contributions (Verissimo and Roberts 2013). Much more important is the fact that a certain quality parameter should be added, thereby showing the importance of reports, otherwise peer review makes no sense (Cintas 2009; Paoletti 2009). There is no obligation at all to review papers. It is better to decline than provide ill-conceived reports, which will be both misleading and useless. Reports should be exhaustive, not necessarily comprehensive, offering constructive criticisms, additional arguments and references, as well as suggestions for improvements and further explorations. Regardless of peer review indexes or referee factors, we all realize that peer review represents a valuable contribution that deserves recognition and should be a part of performance rating (Cintas and Paoletti 2010).

MOVING FORWARDS AND UPWARDS?

If there is a second golden rule of peer review, it is that it requires a community of experts to perform impartial review, accomplished in anonymous form (Hames 2007). In fact, if I may say so, the success of peer review hinges on anonymity; otherwise impartiality cannot be attained.

Inasmuch as reviewers remain anonymous to the authors, the opposite would also be fair, i.e. authors should remain anonymous to the referees. The double-blind review process has been advocated in recent years and top journals like *Science* or *Nature* offer this choice. The underlying principle of double-blind reviewing is that this provides the same opportunities to lesser-known authors as to better-known ones and, in addition, would contribute to remove suspicious bias against minorities, women, some institutions or research arising from third-world countries, among others. However, as noted by some, with the concept of double-blind reviewing we are moving in totally the wrong direction (Zare 2016). This statement appears to be justified because one must remove all identifying data from the submitted manuscript. Even if authors remove personal style and/or self-citation, most referees will still be able to guess who the authors are as research is usually based on previous results.

If one false hypothesis or paradigm is validated by peer review, it could become difficult to refute. Since reviewers are generally unaware of the whole literature or data, by removing all identifications, it will be difficult to assess how trustworthy the results are or where they come from. Surprisingly, a certain subjectivity in reviewer decisions might favor the search for the truth (Park, Peacey and Munafò 2014), thus reducing extrinsic influences that distort the more accurate determination, such as the authorship from a laboratory or previous analyses from other reviewers. This by no means discourages ordinary peer review in favor of double-blind choices.

An alternative to full blindness is a complete and open transparency, as created by the Faculty of 1000 (F1000), a post-publication platform that publishes recommendations in biomedical literature (http://f1000.com/prime), and includes an open peer review system where reviewers' names and their comments are visible on the site (http://f1000research.com). Having said that, however, the concept of transparency requires in addition a clear-cut identification of tools, i.e. how the academic journal (open or not) presents its peer review process to readers and potential contributors (Wicherts 2016).

It is noteworthy to point out another move in scientific publishing: a concluding statement to identify at the end of every paper the individual contributions of all authors. This addendum makes the research process more transparent and adds human dimension to manuscripts (Zare 2016). In principle, this has nothing to do with the peer review process, but the point may play a pivotal role if further retraction takes place and may also contribute to remove false or honorary authorships. This action is welcome, but its analysis and implications should be taken with caution. It is a truism that a research paper is a collective work, even if some use their hands and others their minds. The readership of journals assumes that all individual co-authors are fully aware of the work in question and do accept its interpretation and conclusions at the same extent. Unfortunately, this may not be the actual situation. After detecting invalidating errors, journals and authors alike are reluctant to take actions and retractions may take long. Some co-authors charge others to correct mistakes, claiming they simply wrote the paper or gave advice. Accordingly, we face additional headaches.

Let us consider a well-known example, surely familiar to bioscientists and microbiologists in particular: the Baltimore/Imanishi-Kari case, which was highlighted by mass media (Kevles 1996). In short, Thereza Imanishi-Kari, working as postdoc under the supervision of virologist and Nobel Laureate David Baltimore, co-authored a paper in *Cell* reporting unexpected results on altered genes that regulate immune response in mice. Such results challenged the mechanisms accepted by that time, which would have opened the door to novel treatments against infections like HIV. A researcher at the same lab was unable to reproduce some of the experiments and accused Imanishi-Kari of fabricating the data. Subsequent allegations of misconduct ended up in paper's retraction. During this long and troublesome case, Baltimore defended her against charges, although he had to resign, under pressure, as president of Rockefeller University at NY. A further reinvestigation in 1996, 10 years after publication, exonerated her from misconduct, even though experts found errors in Dr Imanishi-Kari's notebooks and results, but they did not affect the main claims of the paper. Overall, the case served to call for review the integrity of biomedical research. In line with the present discussion herein, this story has a hidden epilogue: the *Cell* paper was co-authored by six researchers, but responsibilities only affected Imanishi-Kari, who presumably executed most of the work, and her supervisor.

PEDAGOGICAL LESSONS

Scientific knowledge is based on reliable and reproducible data. But practitioners of biomedical science know that capricious and erratic behaviors emerge from time to time. Dissenting opinions can bring to light new vistas and prompt researchers to have a further look at evidence and experiments that disagree with their initial hypotheses (Anon. 2016). Peer review should play this important function in science, rather than merely judging the appropriateness of a paper to well-defined guidelines.

Should peer review be taught and learned? Invariably the answer is yes! Learning from peer review, especially by PhD students, will be of enormous benefit to their careers as well as to journals and the scientific community. Introducing peer review as educational element in post-graduate research not only will improve written style and communication skills, but will also provide in-depth insights into the publication process (Le Bailly 2016). In a study conducted at a high-school institution to improve writing, students who reviewed anonymous papers

without receiving peer feedback ('givers') gained more writing ability than those who received solely peer feedback (Lundstrom and Baker 2009). Even retractions offer additional training and wise lessons, although conclusions are so far discouraging: most retractions arise mainly from misconduct and penalties are minimal (Marcus and Oransky 2014). The action of publicizing retractions (e.g. http://retractionwatch.com) sends thoughtful signals to publishers and journals on transparency and how to cut down, if not stop, the vicious circle. In any event, assessment of fabricated or falsified data in classrooms represents a unique educational opportunity for teaching more on data analyses and critical thinking (Contakes 2016).

Improvements in peer review should likely concentrate on post-publication actions, a game involving research teams, editors, as well as journals and publishers. It is convenient for the creation of advanced protocols to identify papers that need further statistical scrutiny, as well as stating clear and unambiguous expressions of concerns, which should be an alert, rather than a priori condemnation (Allison *et al.* 2016).

CONCLUSION

To paraphrase a common aphorism, peer review is the lesser of two evils. It should continue being the gatekeeper to scientific communication, circulating sound experiments, working hypotheses and ideas. This important, time-consuming and daunting task is currently underestimated. Reviewers deserve public recognition by journals, publishers and funding agencies, which should acknowledge by name their invaluable contributions. I would rather advocate the anonymous character of refereeing, which has proven to be useful and fair in most cases. A peer review index may help to boost appreciation. However, its usage and context will also require further attention and analysis. It is hoped that peer review in the 21st century will be improved to make it clearer, transparent and more self-correcting.

Conflict of interest. None declared.

REFERENCES

Allison DB, Brown AW, George BJ *et al.* A tragedy of errors. *Nature* 2016;**530**:27–9.

Anon. Referees' quotes–2010. *Environ Microbiol* 2010;**12**:3303–4.

Anon. Referees' quotes–2010. *Microb Biotechnol* 2011;**4**:108.

Anon. The power of disagreement. *Nat Methods* 2016;**13**:185.

Burnham JC. The evolution of editorial peer review. *J Am Med Assoc* 1990;**263**:1323–9.

Casadevall A, Fang FC. Impact is not importance. *mBio* 2015;**6**:e01593–15.

Cintas P. Increasing visibility and recognition of reviewers. *Els Editors' Update* 2009;**28**:6–7.

Cintas P, Paoletti E. Metrics: include refereeing as part of performance rating. *Nature* 2010;**466**:179.

Contakes SM. Misconduct at the lab? A performance task case study for teaching data analysis and critical thinking. *J Chem Educ* 2016;**93**:314–7.

Fang FC, Casadevall A, Morrison RP. Retracted science and the retraction index. *Infect Immun* 2011;**79**:3855–9.

Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publicateions. *P Natl Acad Sci USA* 2012;**109**:17028–33.

Hames I. *Peer Review and Manuscript Management in Scientific Journals*. Oxford, UK: Blackwell Publishing, 2007.

House of Commons. *Science and Technology Committee, Peer Review in Scientific Publications*, 8th Report of Session 2010–2012. London, UK: The Stationery Office Ltd, 2011.

Ke Q, Ferrara E, Radicchi F *et al.* Defining and identifying sleeping beauties in science. *P Natl Acad Sci USA* 2015;**112**:7426–31.

Kevles DJ. The assault on David Baltimore. *The New Yorker*. New York, NY: Condé Nast, 1996, 94–109.

Kronick DA. Peer review in 18th-century scientific journalism. *J Am Med Assoc* 1990;**263**:1321–2.

Le Bailly B. Learning from peer review. *Nat Nanotechnol* 2016;**11**:204.

Lundstrom K, Baker W. To give is better than to receive: the benefits of peer review to the reviewer's own writing. *J Second Lang Writ* 2009;**18**:30–43.

Marcus A, Oransky I. What studies of retraction tell us. *J Microbiol Biol Educ* 2014;**15**:151–4.

Paoletti E. A reviewer factor (RF) for finding and retaining good reviewers. *Els Editors' Update* 2009;**28**:5–6.

Park IU, Peacey MW, Munafò MR. Modelling the effects of subjective and objective decision making in scientific peer review. *Nature* 2014;**506**:93–6.

Reich ES. *Plastic Fantastic. How the Biggest Fraud in Physics Shook the Scientific World*. New York, NY: Palgrave Macmillan, 2009.

Rousseau R. After the journal impact factor and the web impact factor a referee factor enters the fray: some comments. *ISSI Newsletter*. International Society of Scientometrics and Informetrics, 2006;**2**:2–3.

Scher S. Was Watson and Crick's model truly self-evident. *Nature* 2004;**427**:584.

Silver S. Beyond the fringe: when science moves from innovative to nonsense. *FEMS Microbiol Lett* 2014;**350**:2–8.

Spier R. The history of the peer-review process. *Trends Biotechnol* 2002;**20**:357–8.

Verissimo D, Roberts DL. The academic welfare state: making peer-review count. *Trends Ecol Evol* 2013;**28**:623–4.

Wicherts JM. Peer review quality and transparency of the peer-review process in open access and subscription journals. *PLoS One* 2016;**11**:e0147913.

Wilson R. Referee factor would reward a vital contribution. *Nature* 2006;**441**:812.

Wilson R, Lancaster J. Correction to Referee factor would reward a vital contribution. *Nature* 2006;**441**:1048.

Zare RN. Better practices in scientific publishing. *Angew Chem Int Edit* 2016;**55**:2606–7.