Peer review: rigor? Or rigor mortis?

Vytas A Bankaitis

Recent years have witnessed decreased budgets for scientific research, and these developments sharpen the competition for publication space in high-profile journals. The favorable (and coveted) perceptions that come with such publications bring a decided edge in the competitive arenas of grant funding and career advancements after all. But grave concerns are once again surfacing that this fixation is lending elite journals a distorted and unjustified influence on grant funding, on faculty recruitment, and on career advancement. Such distorted influence is damaging because it miscasts the larger body of impactful and quality science being produced. These concerns are spurring actions within the scientific community ranging from calls for improvements in how scientific research quality and output are evaluated by universities and funding agencies (e.g. DORA; am.ascb.org/dora/), to advocating boycott of elite journals [1], and to the launch of new open-access journals such as *eLife* that advertise themselves as quality-focused alternatives to *Cell*, *Nature*, and *Science*. How effective these measures will be only time will tell, but there is broad agreement that “impact factor” is an abused metric. Much of the problem with the present citation-based impact factor formula lies with it being an absolute-value function that ignores whether citations reference outstanding science, incorrect science, or worst of all, fraudulent publications. It would be an interesting (albeit also imperfect) exercise to re-calculate journal impact factors after positive and negative signs are assigned to their most cited papers.

However we may recoil from a formalized impact factor, the reality is that some manifestation of the impact factor concept is here to stay. This is inevitable so long as there are elite journals. Because the relentless march of scientific progress drives

an exponential increase in scale of the scientific literature, discriminating journals evolve into elitist platforms in the face of such volume. Not all quality science can possibly meet some subjective criterion of “outstanding” after all. So, elitism and impact walk together in lockstep with how a high-profile journal chooses to apply its definition of “outstanding”. This is simply the nature of the beast.

Closely associated with journal exclusivity is a topic of universal concern to practicing scientists and that is the fabric of peer review itself. There is no doubt this process is being scrutinized as never before, and it is here that progress can and must be made. As *EMBO Reports* affords members of its editorial board the opportunity to voice their opinions on matters of general interest, I use this forum to share several thoughts regarding peer review. Various facets of this rather broad topic have been thoroughly discussed in other venues (e.g. [2–4]), so I limit my comments to three considerations: “reviewer experiments”, the question of molecular mechanism and its influence on peer review, and the relationship between peer review and objective assessment of what we truly know.

The “reviewer experiment” problem is discussed by Ploegh [5] and Raff, Johnson and Walter [6]. The seemingly insatiable thirst of reviewers for additional data is enabled by the well-intentioned concept of Supplementary Information, an instrument now run amok which licenses reviewers to demand copious volumes of extra work. Unfortunately, the price exacted by reviewers is not often reflected in dramatic increase in scientific quality, and it often brings debilitating consequences [5,6]. For example, reviewer experiments significantly increase the number of authors on publications, even where most of the work was performed by a few (or one) of the authors.

Author proliferation cannot but raise questions in career evaluations as to who contributed what in the completed study. Reviewer experiments also draw scientists into a difficult territory where significant time and resources are invested in the hope of achieving the prized acceptance of the work for publication in a high-impact journal. Even in cases where the goal is realized, the process frequently extends over a year. Experience tells me the timelines from submission to publication identified by elite journals significantly under-report the actual time it took those bodies of work to navigate review. And these are cases where success was ultimately realized. How often does a laboratory commit to the long road only to be regretfully informed, at its very end, that the revised manuscript still falls short of the advance required for publication in the target journal? Such rejection can come with consequences such as delayed advancement of scientists driving the work and, more painfully, publication by others (during review) of less developed versions of the work in lower-tier journals. Even more troubling, scientific objectivity itself comes under threat of casualty when the elite journal identifies the key result required for publication.

The “tyranny of reviewer experiments” is discussed at length elsewhere [5], but I find grounds for optimism in how editorial policies for peer review are evolving to address this pernicious problem. For example, the EMBO Press journals and others now publish the reviews of all accepted manuscripts. Such transparency presumably suppresses reviewer appetites for demanding unjustifiable amounts of extra work and holds the added benefit of documenting case studies for how papers evolve during peer review. Moreover, although not formalized, the EMBO Press journals encourage cross-referee comments that in effect serve as reviews of the reviews. A related model is

adopted by *eLife*, which tasks assigned reviewers with comparing their evaluations of the submitted work and crafting a consensus response to the editor. Alternatively, rather than commissioning a fourth (or even fifth) reviewer for a submitted manuscript, as frequently occurs in high-profile journals, the handling editor might commission a reviewer or two whose primary charge is to comment on the quality of the scientific reviews. While this particular policy will extend the time frame for initial review, it will ultimately reduce the total time it takes to shepherd a manuscript through the review process. Broad implementation of the approaches taken by the EMBO Press journals and *eLife* help improve the quality and fairness of peer review by better defining how, or whether, a manuscript can be suitably revised. However, once a journal agrees to consider a manuscript, and that manuscript satisfactorily navigates peer review, the work must be published. Editorial rejection of manuscripts at the final stage on the basis of novelty, impact, etc., is unjust. Whereas science magazines have the right to control their “brand” by exercising triage decisions at the level of submission, to do so at the end is wrong.

Why do scientists burden each other with unreasonable reviewer experiments in the first place? One can entertain several scenarios, but I suggest reviewer experiments often result from demands for provision of “molecular mechanism”. The publication bar, particularly in elite journals, is linked to such a standard. But, is the bar for molecular mechanism properly defined? Too often it is not, thereby imposing a disembodied finish line in the name of scientific rigor. Are such demands always appropriate?

“Significant mechanistic advances” assume different meanings in different contexts after all. While it is fair to demand high levels of molecular detail when reporting advances in mature fields, it is unreasonable to set similar bars for discoveries in emerging (i.e. novel) areas of research. So, a poorly defined “molecular mechanism” standard, rather than encouraging scientific rigor, invites collapse into the rigor mortis of reviewer experiments and the drawn-out editorial contortions that accompany them. The bar of molecular mechanism must itself be held to a reasonable experimental description. There is a measure of irony here. Demands for molecular mechanism disproportionately assign impact to already mature areas of research, thereby inviting celebration of incremental progress.

Finally, as Albert Einstein so famously opined, “The greatest obstacle to discovery is not ignorance, but the illusion of knowledge”. The unrestrained obsession of peer review with “significant conceptual advance” and “molecular mechanism” has paradoxically morphed into an effective instrument for promoting illusions of knowledge. As I peruse older scientific literature (1970s/1980s and earlier), I am struck by the measured author discussions of the experimental and interpretive caveats of their work. Those discussions highlighted progress, and provided interesting speculations, while at the same time identifying the remaining questions and uncertainties. In so doing, these bodies of work kept the roads of scientific inquiry open, and inspired young scientists both with ideas and with an appreciation of just how much is unknown. Such objectivity is not a staple of the present literature, and reasonable speculation, even

when identified as such, is muted. I interpret these developments as indications that open objectivity has become a casualty of unfiltered demands for “significant advances in molecular mechanism”. Discussion of open questions and ambiguities in a body of work, whether experimental or interpretive, is discouraged precisely because it invites reviewer experiments. Consequently, uncertainties are massaged and results are spun behind a mask of molecular mechanism. In this way, the literature becomes shrouded with a false sense of progress, and the roads to scientific inquiry close. Peer review would do well to encourage reversal of this unhealthy trend. Such corrective action would promote better practice of the scientific method by delivering a more accurate picture of the limits of knowledge, and by celebrating the core principle that what is known pales in the face of what remains to be discovered.

Acknowledgements

V.A.B. thanks members of the laboratory and his colleagues Todd Graham (Vanderbilt), Chris Burd (Yale), David Katzmman (Mayo Clinic), Paul Herman (Ohio State), Sig Musser (TAMHSC), and David Threadgill (TAMHSC) for critical comments.

References

1. Schekman RW (2013) The Guardian, 9 December.
2. Jefferson T, Wager E, Davidoff F (2002) *JAMA* 287: 2786–2790
3. Smith R (2006) *J Royal Soc Med* 99: 178–182
4. Benos DJ, Bashari E, Chaves JM et al (2007) *Adv Physiol Educ* 31: 145–152
5. Ploegh H (2011) *Nature* 472: 391
6. Raff MC, Johnson A, Walter P (2008) *Science* 321: 36