Peer review is fraught with problems, and we need a fix

November 18 2014, by Andy Tattersall
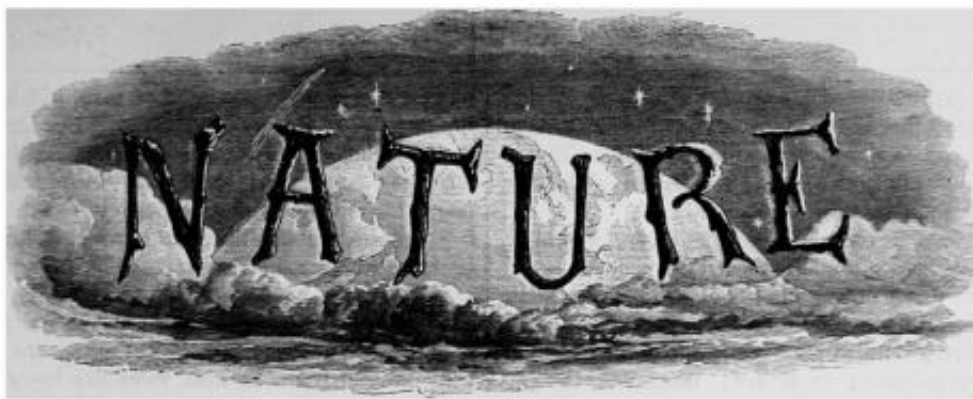
Where it begins. Credit: Nature

Dirty Harry once said, "Opinions are like assholes; everybody has one". Now that the internet has made it easier than ever to share an unsolicited opinion, traditional methods of academic review are beginning to show their age.

We can now leave a public comment on just about anything – including the news, politics, YouTube videos, this article and even the meal we just ate. These comments can sometimes help consumers make more informed choices. In return, companies gain feedback on their products.

The idea was widely championed by Amazon, who have profited enormously from a mechanism which not only shows opinions on a particular product, but also lists items which other users ultimately bought. Comments and star-ratings should not always be taken at face value: Baywatch actor David Hasslehoff's CD "Looking for the Best" currently enjoys 1,027 five-star reviews, but it is hard to believe that the majority of these reviews are sincere. Take for instance this comment from user Sasha Kendricks: "If I could keep time in a bottle, I would use it only to listen to this glistening, steaming pile of wondrous music."

Anonymous online review can have a real and sometimes destructive effect on lives in the real world: a handful of bad Yelp reviews often spell doom for a restaurant or small business. Actively contesting negative or inaccurate reviews can lead to harmful publicity for a business, leaving no way out for business owners.

Academic peer review

Anonymous, independent review has been a core part of the academic research process for years. Prior to publication in any reputable journal, papers are anonymously assessed by the author's peers for originality, correct methodology, and suitability for the journal in question. Peer review is a gatekeeper system that aims to ensure that high-quality papers are published in an appropriate specialist journal. Unlike film and music reviews, academic peer review is supposed to be as objective as possible. While the clarity of writing and communication is an important factor, the novelty, consistency and correctness of the content are paramount, and a paper should not be rejected on the grounds that it is boring to read.

Once published, the quality of any particular piece of research is often measured by citations, that is, the number of times that a paper is formally mentioned in a later piece of published research. In theory, this

aims to highlight how important, useful or interesting a previous piece of work is. More citations are usually better for the author, although that is not always the case.

Take, for instance, Andrew Wakefield's [controversial paper](link) on the association between the MMR jab and autism, published in leading medical journal *The Lancet*. This paper has received nearly two thousand citations – most authors would be thrilled to receive a hundred. However, the quality of Wakefield's research is not at all reflected by this large number. Many of these citations are a product of the storm of controversy surrounding the work, and are contained within papers which are critical of the methods used. Wakefield's research has now been robustly discredited, and the paper was retracted by the *Lancet* in 2010. Nevertheless, this extreme case highlights serious problems with judging a paper or an academic by number of citations.

More sophisticated metrics exist. The [h-index](link), first proposed by physicist Jorge Hirsch, tries to account for both the quality and quantity of a scholar's output in a single number: a researcher who has published n papers, each of which has been cited n times, has an h-index of n. In order to achieve a high h-index, one cannot merely publish a large number of uninteresting papers, or a single extremely significant masterpiece.

The h-index is by no means perfect. For example, it does not capture the work of brilliant fledgling academics with a small number of papers. Recent research has examined a variety of alternative measures of scholarly output, "altmetrics", which use a much wider set of data including article views, downloads, and social media engagement.

Some critics argue that metrics based on tweets and likes might emphasise populist, attention-seeking articles over drier, more rigorous work. Despite this controversy, altmetrics offer real advantages for

academics. They are typically much more fine-grained, providing a rich profile of the demographic who cite a particular piece of work. This system of open online feedback for academic papers is still in its infancy.

Nature journals recently started to [provide](#) authors with feedback on page-views and social media engagement, and sites such as [Scirate](#) allow Reddit-style voting on pre-print articles. However, traditional peer-reviewed journals and associated metrics such as [impact factor](#), which broadly characterises the prestige associated with a particular journal, retain the hard-earned trust of funding organisations, and their power is likely to persist for some time.

Post-publication review

Post-publication review is a model with some potential. The idea is to get academics to review a paper after it has been published. This will remove the bottleneck that journals currently put up because editors are involved and peer review has to be done prior to publication.

But there are limitations. Academics are never short of opinions in their areas of expertise – it goes with the territory. Yet passing comment publicly on other people's research can be risky, and negative feedback could provoke a retaliation.

Post-publication review also has the potential for bias via preconceived judgements. One researcher may leave harsh comments on another's research based on the fact they do not like that person: rivalry in academia is not uncommon. Trolling on the web has become a serious problem in recent times, and it is not just the domain of the uneducated, bitter and twisted section, but is also enjoyed by members of society who are supposedly balanced, measured and intelligent.

One post-publication review platform, [PubPeer](#), allows anonymous commenting – which, as seen with sites that allow for anonymous posts – could open the door for more trolling and abusive behaviour. It could offer reviewers an extra level of protection from what they say. One researcher recently filed a lawsuit over anonymous comments on PubPeer which they claim caused them to lose their job, after accusations of misconduct in their research. In a similar case, an academic claimed to have lost project funding after a reviewer complained about a blog post they had written about their project.

Post-publication comment can also be susceptible to manipulation and bias if not properly moderated. Even then, it is not easy to detect how honest and sincere someone is being over the Web. Recent stories featuring TripAdivisor and the independent health feedback website Patient Opinion show how rating and review systems can come into question. Nevertheless, research can possibly learn something from the likes of Amazon in how a long tail of research discoverability can be created. Comments and reviews may not always truly highlight how good a piece of research is, but they can help create a post-publication dialogue, a connectivity, globally about that topic of research, that in time sparks new ideas and publications.

Many now believe that long-standing metrics of academic research – peer review, citation-counting, impact factor – are reaching breaking point. But we are not yet in a position to place complete trust in the alternatives – altmetrics, [open science](#), and post-publication review. But what is clear is that in order to measure the value of new measures of value, we need to try them out at scale.

This story is published courtesy of [The Conversation](#) *(under Creative Commons-Attribution/No derivatives).*