SCIENTIFIC METHOD —

Journal tries crowdsourcing peer reviews, sees excellent results

Approach allows reviewers to focus on the stuff they know best, speeds up process.

CHRIS LEE (US) - 8/6/2017, 21:03

Peer Review
Arbitraje
Revisión por expertos

Enlarge

Peer review is supposed to act as a sanity check on science. A few learned scientists take a look at your work, and if it withstands their objective and entirely neutral scrutiny, a journal will happily publish your work. As those links indicate, however, there are some issues with peer review as it is currently practiced. Recently, Benjamin List, a researcher and journal editor in Germany, and his graduate assistant, Denis Höfler, have come up with a genius idea for improving matters: something called selected crowd-sourced peer review.

A cynic's view of peer review

Pardon a bit of editorializing here. The cynical and somewhat accurate view of peer review is that a few grumpy old men are given prior access to your work. Depending on their mood, medication, and memory, one of three things happens: they forget about it, leaving you and your manuscript languishing in limbo—remember that scientific productivity is measured in papers, so delayed papers means no productivity.

Or, the reviewers may happen to remember your paper but couldn't really give the tiniest hint of caring about your work, so they write some bland and unhelpful comments that seem essentially random. Most entertaining of all, though, is the researcher who discovers that you've just done what they really wanted to do, and they would rather be damned to an eternity of committee work than let you publish. (Of course, I should note that I fall into none of these categories, and my reviews are a picture of perfect peer performance.)

The utter randomness of peer review is frustrating for everyone. Papers get delayed, editors get frustrated, the responsible reviewers get overloaded. Even when everyone is trying their best, any set of three reviewers can disagree so wholeheartedly about your work that the editor has to actually think about a decision—something no editor ever wants to be faced with. (I should also note here that every editor that dealt with my work has always made thoughtful and logical decisions.)

Yet another makeover

My central point: peer review is burdensome and sometimes barely functional. So how do we improve it? The main way is to experiment with different approaches to the reviewing process, which many journals have tried, albeit with limited success. Post-publication peer review, when scientists look over papers after they've been published, is also an option but depends on community engagement.

But if your paper is uninteresting, no one will comment on it after it is published. Pre-publication peer review is the only moment where we can be certain that someone will read the paper.

So, List (an editor for Synlett) and Höfler recruited 100 referees. For their trial, a forum-style commenting system was set up that allowed referees to comment anonymously on submitted papers but also on each other's comments as well. To provide a comparison, the papers that went through this process also went through the traditional peer review process. The authors and editors compared comments and (subjectively) evaluated the pros and cons. The 100-person crowd of researchers was deemed the more effective of the two.

The editors found that it took a bit more time to read and collate all the comments into a reviewers' report. But it was still faster, which the authors loved. Typically, it took the crowd just a few days to complete their review, which compares very nicely to the usual four to six weeks of the traditional route (I've had papers languish for six months in peer review). And, perhaps most important, the responses were more substantive and useful compared to the typical two-to-four-person review.

So far, List has not published the trial results formally. Despite that, Synlett is moving to the new system for all its papers.

Why does crowdsourcing work?

Here we get back to something more editorial. I'd suggest that there is a physical analog to traditional peer review, called noise. Noise is not just a constant background that must be overcome. Noise is also generated by the very process that creates a signal. The difference is how the amplitude of noise grows compared to the amplitude of signal. For very low-amplitude signals, all you measure is noise, while for very high-intensity signals, the noise is vanishingly small compared to the signal, even though it's huge compared to the noise of the low-amplitude signal.

Our esteemed peers, I would argue, are somewhat random in their response, but weighted toward objectivity. Using this inappropriate physics model, a review conducted by four reviewers can be expected (on average) to contain two responses that are, basically, noise. By contrast, a review by 100 reviewers may only have 10 responses that are noise. Overall, a substantial improvement. So, adding the responses of a large number of peers together should produce a better picture of a scientific paper's strengths and weaknesses.

Didn't I just say that reviewers are overloaded? Doesn't it seem that this will make the problem worse?

Well, no, as it turns out. When this approach was tested (with consent) on papers submitted to Synlett, it was discovered that review times went way down—from weeks to days. And authors reported getting more useful comments from their reviewers.

The forum is open, so as a reviewer, you can see comments accumulating, and you know the editor is going to close comments at some point soon. You are either going to do the job now or not do it at all—you can't put the editor off for three weeks before deciding that you don't have time.

I think there is also a psychological aspect to crowd reviewing. If you know that you are only one of two to four people reviewing a manuscript, and you know from experience that 75 percent of reviews are crap (the ones done by the other three, obviously), then you feel that you have to really *work* at your review.

Time sinks

This work's time-consuming aspects comes in several flavors: first, there are very few papers that are entirely in my field(s) of expertise. That means that I can blast through the bits I know well and, after some thought, hopefully offer a bit of insightful commentary on the author's results. But, then I hit the bits that aren't in my wheelhouse. I have to slow down and try to understand them as well. The result is usually a mixture of relevant commentary and random expectorations that demonstrate my own ignorance. (Put differently, parts are a waste of everyone's time.)

To give you a quick example: if an author makes a statement in the introduction, they will support that statement with a reference to an older paper. If I am absolutely confident with the material, I will be familiar with both the statement and the reference and can quickly move on. What if the statement is just on the edge of my field, or the reference is unfamiliar? Then I have to make a decision: do I do my job properly and follow the reference, or do I trust that the editor has picked another referee who knows that material better than I do?

With a crowd of reviewers, I can confidently focus on the parts that I am most expert in and only comment on that. And, since I can see the comments of the other reviewers, I don't have to repeat what the others have said or delve into areas that I am unfamiliar with. If a reviewer has misunderstood some aspect, I can add a comment to their comment, so the review process becomes a discussion.

You're not always going to save time, however. From personal experience, I know I've spent hours re-deriving results and doing order-of-magnitude calculations to check data and theory. This still has to be done and won't get any faster, the difference being that I can do only that part and leave that as my contribution to the review.

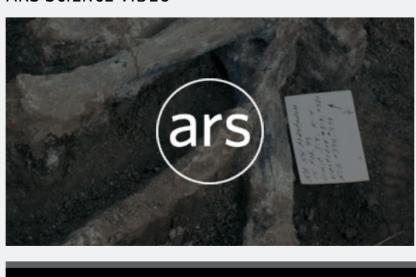As a result, and as List reported, the review will be as detailed (and probably more so) than that obtained from just a few reviewers. Bad papers will die a death of a thousand cuts, while good papers stand a reasonable chance of surviving the review process. For authors, the outsized influence of a single grumpy reviewer is reduced, much to everyone's relief.

There are big questions left. Will it scale? Is the effectiveness that List has seen simply enthusiasm for something new? If so, the anonymous forum may become a ghost town, forcing editors back to having to nag reviewers to respond. The question of scale is more difficult. In some ways, I think this can work because it distributes the load more evenly across scientists. But I'm not overly confident. Many alternatives to peer review have been tried, and none have yet caught on.

This post originated on Ars Technica

ARS SCIENCE VIDEO >

Incredible discovery places humans in California 130,000 years ago

CHRIS LEE
Chris writes for Ars Technica's science section. A physicist by day and science writer by night, he specializes in quantum physics and optics. He is delocalised, living and working in Eindhoven and Enschede, the Netherlands.
EMAIL chris.lee@arstechnica.com   TWITTER @exMamaku

READER COMMENTS   35

SHARE THIS STORY

← PREVIOUS STORY

Related Stories

Today on Ars