Crowd-based peer review can be good and fast

Confidential feedback from many interacting reviewers can help editors make better, quicker decisions, explains **Benjamin List**.

When it works — and that's much of the time — peer review is a wondrous thing. But all too often, it can be an exercise in frustration for all concerned. Authors are on tenterhooks to learn of potentially career-changing decisions. Generous peer-reviewers are overwhelmed. And editors are condemned to doggedly sending reminders weeks after deadlines pass. When the evaluation finally arrives, it might be biased, inaccurate or otherwise devoid of insight. As an author and latterly editor-in-chief of *Synlett*, a chemical-synthesis journal, I've seen too many 'reviews' that say little more than "this manuscript is excellent and should be published" or "this manuscript clearly doesn't reach the standards for your journal".

So last year, my graduate student and editorial assistant Denis Höfler and I started to test an alternative approach, which I call intelligent crowd reviewing. Now we plan to use it as our main tool for evaluating manuscripts. Why, we reason, in a world of online bulletin boards and multi-authored encyclopedia entries, should we conduct peer review in much the same way as when manuscripts were delivered by postal workers with horses?

I am not proposing what is sometimes referred to as crowdsourced reviewing, in which anyone can comment on an openly posted manuscript. I believe that anonymous feedback is more candid, and that confidential submissions give authors space to decide how to revise and publish their work. I envisioned instead a protected platform whereby many expert reviewers could read and comment on submissions, as well as on fellow reviewers' comments. This, I reasoned, would lead to faster, more-informed editorial decisions.

When I began discussing the idea with colleagues in Germany's Max Planck Society and at *Synlett*, they were rather sceptical. Fellow editors worried that they might be flooded with responses and would miss out on the perfect expert, the one referee whose opinion was definitive. Others feared that 'power-referees' would dominate, that confidentiality might be breached, and that such a project would not be scalable: that if too many journals tried this, they would cannibalize each other's reviewers.

As an experimentalist, I decided to test the theory. With Denis, I recruited just over 100 highly qualified referees, mostly suggested by our editorial board. We worked with an IT start-up company to create a closed online forum and sought authors' permission to have their submissions assessed in this way. Conventional peer reviewers evaluated the same manuscripts in parallel. After an editorial decision was made, authors received reports both from the crowd discussion and from the conventional reviewers.

In May last year, we began to upload manuscripts on to the platform one at a time, and were impressed with the overwhelming number of responses collected after only a few days. This January,

> ## WE PUT UP TWO MANUSCRIPTS SIMULTANEOUSLY AND GAVE THE CROWD 72 HOURS TO RESPOND.

we put up two manuscripts simultaneously and gave the crowd 72 hours to respond. Each paper received dozens of comments that our editors considered informative. Taken together, responses from the crowd showed at least as much attention to fine details, including supporting information outside the main article, as did those from conventional reviewers.

So far, we have tried crowd reviewing with ten manuscripts. In all cases, the response was more than enough to enable a fair and rapid editorial decision. Compared with our control experiments, we found that the crowd was much faster (days versus months), and collectively provided more-comprehensive feedback.

Our authors reacted positively, saying that they appreciated the comprehensiveness of the crowd's comments and the speedy turnaround. The authors of the one manuscript that was rejected did not complain. As editors we did have more to read, but did not feel that our workload was massively increased: a crowd report is typically no harder to digest than three or four conventional reviews.

So far, no referees have been domineering in the discussions, although if referees were to act questionably in general, it would be straightforward to replace them, because editors know who referees are. Similarly, we found no breaches of confidentiality in this limited experiment.

Can this approach be sustained, or even expanded? In our system, referees do potentially see more manuscripts than in conventional peer reviewing, but they comment only on papers they choose to read, and only as much as they are inclined to do. They also seem to enjoy interacting with others, rather than writing alone. Nonetheless, we are considering ways to acknowledge our most reliable referees, perhaps with free subscriptions or by naming them on our website. Granted, assembling a suitable crowd is easier for a specialized journal catering to a relatively small community. In our case, all participants are interested in chemical synthesis. In the future, we plan to use key words to automatically match manuscripts with suitable referees, a strategy that would apply for broader journals.

We are in the early stages of switching to crowd reviewing as our main tool for evaluating manuscripts, and we expect that there will be challenges ahead. We plan to maintain a highly functioning crowd, and our efforts will include learning how to keep reviewers engaged once the novelty has worn off. The experiment continues, but my conclusion so far is clear: crowd reviewing works. ■

Benjamin List *is a director at the Max Planck Institut für Kohlenforschung in Mülheim an der Ruhr, Germany, and the editor-in-chief of* Synlett.
e-mail: list@mpi-muelheim.mpg.de