



Peer reviewers need more nurturing

Richard Catlow explains why three national scientific societies are calling for research evaluators to be highly valued and trained.

When peer review is broken, so is science. That is why this month, three national scientific academies — the French Academy of Sciences, the German Leopoldina and the UK Royal Society — are issuing a joint statement on how to make sure research evaluation is done well (see go.nature.com/2ayvlsf). This is the first time the societies have spoken out on the issue, and they do so at the behest of Carlos Moedas, the European Union commissioner for research, science and innovation. As foreign secretary of the Royal Society, I helped to put this statement together. I hope it will influence all involved in assessing scientists for promotion, tenure and awards.

Our key recommendation is that peer review should remain the cornerstone of assessment. It must be carried out by people who are competent peers — and who are recognized as such. These reviewers

need the time and training to examine scientific contributions thoughtfully, without depending on bibliometric summaries. To make that happen, we must treat assessment expertise as a valuable resource.

If assessment is to work well, it is important not to over-assess. Too much time is spent reviewing and re-reviewing. In the United Kingdom, for example, we have seen a growth of mid-term reviews for large projects. Of course we need to check that such projects stay on track, but intense reviews drain the time of senior scientists, making them less available for more important assessments. We need to have more confidence in the people chosen to lead big projects. Modest checks can assess whether the work is proceeding well and trigger more-thorough investigations when there are actually signs of problems.

Confidence in reviewers is also important. Those selected to conduct peer review should be esteemed by the individuals and communities they assess. Relevant expertise should be the lead criterion — scientists should feel that their proposals or contributions were assessed accurately because reviewers came from the right discipline. And the experts called in to perform reviews cannot be a closed club, whose members could be inclined to choose people like themselves. Excellence is the primary qualification, so gender, race, ethnicity, disability, sexual orientation, age and so on must be no barrier to inclusion on a panel of assessors. Increasing the number of people who are asked to review will also ensure that reviewers are not stretched too thin.

Societies and institutions should take measures to build reviewing expertise. Our current system presumes that scientists will simply pick up the necessary skills. But reviewers need to be trained in how to think about conflicting input from referees and how to compare projects and proposals across different fields. They should also be warned against too heavily rewarding topics that are currently fashionable. They should be taught about unconscious bias and techniques to

guard against it. Some programmes for this are already in place: the UK Engineering and Physical Sciences Research Council, for example, has run mock panels and asked novice reviewers to evaluate projects using proposals and information from previous panels.

Following these recommendations would prepare us to tackle what in my view is the most worrying aspect of research evaluation: the over-reliance on metrics. This distorts the research programmes of early-career scientists. I have seen younger colleagues, in what should be a highly creative stage of their careers, slant their research towards topics they believe will accrue large numbers of citations and appear in journals with high impact factors. Evidence suggests important questions are neglected as a result.

I was lucky. When I began my scientific career in the 1970s, I

had no real sense of how my work was cited. My discipline — computational materials-chemistry — was barely acknowledged by main-stream chemists. If I had been citation-driven, I might have abandoned a field that is now central to developing sophisticated materials including porous catalysts, electronic ceramics and ionically conductive materials. By the 1990s, when citation data became prominent, I was already a full professor.

Metrics cannot be a proxy for expertise. I chaired the chemistry panel of the 2014 Research Excellence Framework, which assessed research units at UK universities. Reviewers read the actual papers, as well as looking at citation data. Bibliometrics should be only one strand of evidence.

Similarly, impact factors tell us about a journal; they cannot be used as a measure of the quality of an individual article in that journal. It was five years ago this month that members of the scientific community launched what is now known as the San Francisco Declaration on Research Assessment, arguing against the use of journal-based metrics to stand in for the quality of individual scientists. Almost 900 organizations have signed on, yet actual changes in behaviour have been slow in coming. I have seen recent cases in which applicants for promotion were obliged by their university to give the impact factors of the journals they had published in.

Overstretched and insecure reviewers reach for bibliometrics because they are easy and quantitative. The impetus for change will not come from ever more arguments against them, but from freeing up and creating more human capacity for research assessment. I hope this month's Three Academy Statement will encourage academic leaders and scientific funders to do so. ■

EXPERTS CALLED
IN TO
PERFORM
REVIEWS
CANNOT
BE A
CLOSED CLUB.

Richard Catlow is a professor of materials chemistry and catalysis at University College London and Cardiff University, UK. e-mail: richard.catlow@royalsociety.org