



Take peer pressure out of peer review

Until we study the social dynamics of review panels, assessments will be suboptimal, explains Gemma Derrick.

In its most recent round of university assessments, the United Kingdom tried something new. To judge the value of research beyond academia, review panels in 2014 included a much greater proportion of non-academics than in the previous assessment. For example, pharmaceutical scientists evaluated output in clinical medicine, and government infrastructure experts sat on engineering panels. Despite the shake-up, the university rankings changed little.

That was probably not because the experts within and outside academia agreed on what makes for research that has real-world value. Instead, it seems that the non-academics had little influence.

I interviewed reviewers before and after the evaluations — I study the culture of knowledge production. Those of all backgrounds told me that the interactions on panels were very much academically led. The non-academics had trouble penetrating what one described as “quite a strong culture”. Academics acknowledged and accepted that outsiders were sidelined: their value was in validating the evaluation. One called their presence “a bit of tokenism”; another said that it “provided a type of political capital”.

The UK education council is now assembling reviewers for its next assessment in 2021. It is mainly concerned with getting a mix of academics and non-academics onto panels, and selecting which industries to represent. Without a strategy for how members will work together, these are meaningless efforts.

I think that reviewers of different stripes are not genuinely reaching consensus. There are too few practices to help them do so, and too little knowledge is available to develop these tools. Much work has been done on how to get experts to come to better decisions, but it is unclear how well it applies to confidential reviewing panels.

If peer review is to work as intended (and as commonly assumed), we need to make sure that diverse perspectives are considered amid consistently cliquy groups of academics. In other words, before funding agencies shove a group of strangers into a room and insist they deliver a decision within a strict time limit, we need a better understanding of how these panels actually function.

My own and others’ observations show that a peer-review panel is not like some collaborative mural, where everyone contributes a piece to the picture. It is more like a tug of war — with a rope that has many ends. Evaluators form alliances and join various ends of the rope. This sets the panel’s dominant mode for dictating how all proposals are assessed. Those outside this framework are quickly silenced, even if they were recruited for their perspective.

The situation undermines what peer review is supposed to accomplish. Peer review is esteemed because, unlike assessments based on metrics, it can incorporate human judgement: panellists

are charged with considering how well a project fits particular goals or with accounting for mitigating circumstances, such as illness, in researchers’ productivity.

The system rests on the assumption that experts will work together to air, debate and consider varied views. A sustained collective effort is expected to manage conflicts, to catch weaknesses and mistakes, and to make sound judgements about how to spend public money. These presumed qualities provide political legitimacy, and the outcomes of academic evaluation are accepted by the wider community.

These benefits accrue only if the process is perceived to be fair and informed. Meanwhile, troublingly reductive metric-based evaluations threaten to dominate, with performance defined by strictly measurable formulae. These cost less and can be touted as more objective.

I think the better investment is in learning how to evaluate and improve human review. Unfortunately, efforts to assess review are often thwarted. Most studies of panels can consider only the inputs and outputs, without understanding why some proposals find favour, but not others. Because confidentiality is so prized, getting access to panels took me more than a year. This is at odds with the drive for science and for government decision-making to become more accountable. Worse, it stifles efforts to improve. The study of review panels is essential to optimize the process and to demonstrate that optimal review is valued.

Even limited observations are yielding preliminary pointers that can themselves be evaluated. For example, the Swedish Research Council recently suggested that assigned seating could keep panel members who are already well known to each other from sitting together, and so encourage participation by women and international members.

Many other ideas are worth trying. Pre-evaluation training could help panellists understand how, consciously or otherwise, they might silence competing ideas. Splitting panels into experts who evaluate proposals in isolation and others who make decisions based on blinded assessments could reduce groupthink. Last, peer-review panels could include a non-academic chair to encourage debate from all and actively challenge the consensus.

We should test these strategies with quasi-experimental simulations and by directly observing more panels in action. To ensure the future of peer review, we must understand how to do it better. ■

A PEER-REVIEW
PANEL IS NOT
LIKE SOME
**COLLABORATIVE
MURAL;**
IT IS MORE LIKE A
TUG OF WAR.

Gemma Derrick co-directs the Centre for Higher Education, Research and Evaluation at Lancaster University, UK. Her book, *The Evaluators’ Eye* is published this month.
e-mail: g.derrick@lancaster.ac.uk