A personal take on science and society

World view

Pandemic researchers – recruit your own best critics

By Daniël Lakens

To guard against rushed and sloppy science, build pressure testing into your research.

A s researchers rush to find the best ways to quell the COVID-19 crisis, they want to get results out ultra-fast. Preprints – public but unvetted studies – are getting lots of attention. But even their advocates are seeing a problem. To keep up the speed of research and reduce sloppiness, scientists must find ways to build criticism into the process.

Finding ways to prove ourselves wrong is a scientific ideal, but it is rarely scientific practice. Openness to critique is nowhere near as widespread as researchers like to think. Scientists rarely implement procedures to receive and incorporate pushback. Most formal mechanisms are tied to the peer-review and publishing system. With preprints, the boldest peers will still criticize the work, but only after mistakes are made and, often, widely disseminated.

An initial version of a preprint by researchers at Stanford University in California estimated that COVID-19's fatality rate was 0.12–0.2% (E. Bendavid *et al.* Preprint at medRxiv http://doi.org/dskd; 2020). This low estimate was removed from a subsequent version, but it had already received widespread attention and news coverage. Many immediately pointed out flaws in how the sample was obtained and the statistics were calculated. Everyone would have benefited if the team had received this criticism before the data were collected and the results were shared.

It is time to adopt a 'red team' approach in science that integrates criticism into each step of the research process. A red team is a designated 'devil's advocate' charged with finding holes and errors in ongoing work and to challenge dominant assumptions, with the goal of improving project quality. The team has a role similar to that of 'white-hat hackers' hired in the software industry to identify security flaws before they can be discovered and exploited by malefactors. Similarly, teams of scientists should engage with red teams at each phase of a research project and incorporate their criticism. The logic is similar to the Registered Report publication system – in which protocols are reviewed before the results are known – except that criticism is not organized by journals. Ideally, there is a larger amount of speedier communication between researchers and their red team than peer review allows, resulting in higher-quality preprints and submissions for publication.

Even scientists who invite criticism from a red team acknowledge that it is difficult not to become defensive. The best time for scrutiny is before you have fallen in love with your results. And the more important the claims, the more scrutiny they deserve. The scientific process needs to incorporate methods to include 'severe' tests that will

> **"Researchers need to commit to addressing criticism from the outset."**

Daniël Lakens is an associate professor in the human–technology interaction group at Eindhoven University of Technology, the Netherlands. e-mail: d.lakens@ tue.nl

prove us wrong when we really are wrong.

An example of a large-scale collaboration that applies a red-team approach is the Psychological Science Accelerator (PSA), a global network of more than 500 psychology laboratories. The PSA has solicited research projects on questions related to the COVID-19 pandemic and has offered to assist with data collection. Projects range from effective risk communication to cognitive-reappraisal interventions. After researchers develop protocols, the PSA assembles a red team of experts in research ethics, measurement, data analysis and the project's field to offer criticism and to allow researchers to revise their protocols.

I reviewed one of these protocols after it had been submitted to a journal. I later saw the PSA reviews and learnt that I had repeated many criticisms, such as the generalizability of the stimulus and flexibility of the data analysis, that the red team had made – and that the researchers had opted to ignore.

This shows that assembling a red team isn't enough: research teams need to commit to addressing criticism from the outset. Sometimes, this is straightforward – items on checklists are absent from a proposal, or an independent statistical analysis yields different results, for example. Usually, it will be less clear whether criticism merits changing a protocol or including a caveat. The key is that, when results are presented, the team transparently communicates the criticism that the red team raised. (Perhaps incorporated criticism could be listed in the methods section of a paper, and unincorporated criticism in the limitations.) This will show how severely a claim has been tested.

Pushback on each step of a project should be recognized as valuable quality control and adherence to scientific values. Ideally, a research team could recruit its red team from group members not immediately involved in the project.

Incentives for red teams in science deserve special consideration. A red team might identify major flaws that mean a study should not proceed, so including a team member as a co-author on a future publication by the group would be a conflict of interest. In the computer-security industry, a red team is often paid if it uncovers serious errors. Computer scientist Donald Knuth famously gave out 'bug bounties' to people who uncovered technical errors in his published work. (Recipients often kept the small cheques as souvenirs, suggesting that social credit works as an incentive.) To investigate incentivized criticism, my group is now recruiting red-team members and offering financial rewards (see go.nature.com/3frpbjq).

With research moving faster than ever, scientists should invest in reducing their own bias and allowing others to transparently evaluate how much pushback their ideas have been subjected to. A scientific claim is as reliable as only the most severe criticism it has been able to withstand.