THIS WEEK

EDITORIALS

DRUGS New agreement to tackle pharmaceutical pollution **p.164**

WORLD VIEW Vaccination the best way to measure health care **p.165**

DUNG OVER Rolling beetles fooled by look-alike seeds **p.167**

Let's think about cognitive bias

The human brain's habit of finding what it wants to find is a key problem for research. Establishing robust methods to avoid such bias will make results more reproducible.

"Ever since I first learned about confirmation bias I've been seeing it everywhere." So said British author and broadcaster Jon Ronson in *So You've Been Publicly Shamed* (Picador, 2015).

You will see a lot of cognitive bias in this week's *Nature*. In a series of articles, we examine the impact that bias can have on research, and the best ways to identify and tackle it. One enemy of robust science is our humanity — our appetite for being right, and our tendency to find patterns in noise, to see supporting evidence for what we already believe is true, and to ignore the facts that do not fit.

The sources and types of such cognitive bias — and the fallacies they produce — are becoming more widely appreciated. Some of the problems are as old as science itself, and some are new: the IKEA effect, for example, describes a cognitive bias among consumers who place artificially high value on products that they have built themselves. Another common fallacy in research is the Texas sharp-shooter effect — firing off a few rounds and then drawing a bull's eye around the bullet holes. And then there is asymmetrical attention: carefully debugging analyses and debunking data that counter a favoured hypothesis, while letting evidence in favour of the hypothesis slide by unexamined.

Such fallacies sound obvious and easy to avoid. It is easy to think that they only affect other people. In fact, they fall naturally into investigators' blind spots (see page 182).

Advocates of robust science have repeatedly warned against cognitive habits that can lead to error. Although such awareness is essential, it is insufficient. The scientific community needs concrete guidance on how to manage its all-too-human biases and avoid the errors they cause.

That need is particularly acute in statistical data analysis, where some of the best-established methods were developed in a time before data sets were measured in terabytes, and where choices between techniques offer abundant opportunity for errors. Proteomics and genomics, for example, crunch millions of data points at once, over thousands of gene or protein variants. Early work was plagued by false positives, before the spread of techniques that could account for the myriad hypotheses that such a data-rich environment could generate.

Although problems persist, these fields serve as examples of communities learning to recognize and curb their mistakes. Another example is the venerable practice of double-blind studies. But more effort is needed, particularly in what some have called evidence-based data analysis: research on what techniques work best to establish default analytical pipelines for cleaning and debugging data sets, selecting models and other steps of analysis.

More specifically, science needs ways to identify the mistakes most likely to be made by novice (and not-so-novice) number crunchers. The scientific community must design research protocols that safeguard against these errors, and devise methods that ferret out sloppy analyses.

NATURE.COM
For *Nature*'s special collection on reproducibility, see:
go.nature.com/huhbyr

Some researchers already do this well, so one relatively simple strategy is to improve how knowledge and resources move from a narrow group of experts to the broader scientific community. If highly respected, easy-to-implement alternative routes are available and encouraged, it will be harder to cling to analyses that are rigged by conscious or unconscious bias to produce the results that researchers want. Funders should support teams that are attempting to determine the best analytical routes, and should provide training in data analysis for others. Institutions and principal investigators should make such training mandatory.

Finally, the scientific community must go beyond statistical safeguards, and improve researchers' behaviour. Angst over unreliable research has already spurred investigations into ways to make results more robust. Some of the most promising address not just techniques, but also academic culture: laboratory and workplace habits can discourage rigour, or can enforce it through blinding, preregistering analytical plans, crowdsourcing analysis, formally laying out null and alternative hypotheses, and labelling analyses as exploratory or confirmatory.

Such strategies require effort, but offer significant rewards. Blind analysis forces creative thinking as researchers struggle to find explanations for hypothetical results. A Comment on page 187 explores these rewards and offers tips for researchers ready to try it.

"It is easy to think that fallacies only affect other people."

Crowdsourcing shows how the same data set, analysed with different approaches, can yield a variety of answers; it is a reminder that single-team analysis is only part of the story. As a Comment on page 189 reveals, crowdsourced analyses and interdisciplinary projects can also compare analysis techniques across disciplines, and show how one field might hold lessons for another. Some differences in approach are probably down to cultural happenstance — "we have always done it this way" — rather than to selection of best practice. That should change.

To ensure that such practices actually strengthen science, scientists must subject the strategies themselves to scientific scrutiny. (No one should take recommendations to counter bias on faith!) Social scientists have an important role here — studies of science in action are essential. Careful observation of scientists can test which strategies are most effective under what circumstances, and can explore how debiasing strategies can best be integrated into routine scientific practice.

Funders should support efforts to establish the best methods of blind analysis, crowdsourcing and reviewing registered analysis plans, and should help meta-scientists to test and compare these practices. Ideally, the utility and burdens of these strategies under varying circumstances would be explored and published in the peer-reviewed literature. This information could then be fed into much-needed training programmes, and so better equip the next generation of scientists to do good science.

Finding the best ways to keep scientists from fooling themselves has so far been mainly an art form and an ideal. The time has come to make it a science. We need to see it everywhere. ■

© 2015 Macmillan Publishers Limited. All rights reserved