

# World view



By Nicole C. Nelson

## Understand the real reasons reproducibility reform fails

**Lack of rigour is often blamed on pressure to publish. But ethnographers can find out what truly keeps science from upping its game.**

A decade ago, the US National Institute of Neurological Disorders and Stroke convened a workshop on how to improve the rigour of preclinical research. Its recommendations were surprisingly straightforward: scientists should mask (or ‘blind’) their studies; randomize; estimate appropriate sample sizes; and specify rules for data handling (S. C. Landis *et al. Nature* **490**, 187–191; 2012). Ten years on, many preclinical scientists still do not take these basic steps.

Ask most advocates of rigorous science why this is, and they will answer with two words: perverse incentives. Scientists are rewarded for getting things published, not for getting things right, and so they tend to favour speed and ease over robustness. But as an ethnographer, this explanation has never sat well with me. I’ve spent more than 15 years studying biomedical research cultures, and scientists’ behaviours are rarely so transactional. So I decided to knock on a few doors at my institution, the University of Wisconsin–Madison, to ask researchers who work with animals why they were using the methods they were.

They explained that their decisions were based not on publication pressures, but on maintaining the integrity of experiments and respecting facility routines. For example, techniques to mask which treatment groups animals are in risk causing misidentification or cross-infection.

One scientist told me how years of tedious mouse breeding had gone off the rails when animals’ ear tags fell out, cage cards were swapped or spreadsheets had errors. The risk of mix-ups caused by masking seemed too great, especially in experiments where effects were so pronounced that there was no risk of bias. “You’d have to be fooling yourself pretty hard to see tumours where there are none,” he laughed. Yet he did think masking was important if the effects of an intervention were small. As we spoke, he realized that he had been so focused on avoiding the ‘horrors of misidentification’ that he might have dismissed other risks. In his case, one set of fears crowded out another.

Labour structures also make masking fraught. An investigator in a collaborative clinical project can pass tissue samples to a pathologist without explaining the sample numbering system. But in preclinical research, lab members are expected to take projects from beginning to end, providing few natural opportunities for masking. One graduate student didn’t want to ‘beg’ for labour from fellow students to allow her to mask her studies. She also worried that colleagues might be less conscientious about tasks that were important for her progress but not their projects.

**Challenging experimental designs is best done after asking what’s behind them.”**

**Nicole C. Nelson** is an associate professor in the Department of Medical History and Bioethics at the University of Wisconsin–Madison. e-mail: nicole.nelson@wisc.edu

Efforts to implement masking often run counter to policies at animal facilities. US guidelines state that genotype information should be listed on cage cards, which effectively unmask many studies. Other policies pose similar conundrums. Segregating mice treated with a viral vector prevents cross-contamination, but also reveals the treatment group. Many scientists were reluctant to ask facilities to alter their routines or make exceptions.

The main considerations I found for whether to mask were integrity, collegiality and animal welfare. Dig deeper on failures to adopt other reproducibility reforms, and I’m confident that you’ll find more than perverse incentives.

Ethnography excels as a tool for disrupting oversimplified stories about decisions. Sociological expertise was key to understanding how the culture at NASA led to the 1986 *Challenger* disaster, in which a Space Shuttle exploded shortly after launch. More recent ethnographic work at the agency has shown how the administrative structure of a mission affects the science it produces. This work enables more-informed choices on research culture and other issues.

I’ve experienced this first-hand. As the embedded ethnographer in an early cancer-genomics trial, I helped the team to understand problems in implementing this new type of study. These came down to conflicting beliefs about what constitutes high-quality data. Bench scientists wanted regular technology upgrades to get ever more accurate genomic information; clinical researchers wanted set protocols that rarely changed. Understanding this conflict helped the team to work through it.

Unearthing assumptions, fears and social relations does not have to mean leaving scientific practice unchallenged. In the 1990s, scientists assumed that the effect of knocking out a gene in an organism would be so pronounced that there was no need for controls – until failures to replicate started popping up. But challenging experimental designs is best done after asking what’s behind them. And it’s worth considering that, in some situations, the net benefit of making practices more rigorous might be minimal.

Scientific values can drive change. Reproducible workflows can be implemented in ways that assuage fears rather than stoking them. One scientist had an elaborate system of masking, involving colour-coded sample tubes and cage cards, and often had to fight the animal core facility to use it. “I just know that we’re human,” she told me, “and humans are always going to be biased.” Her commitment to an ideal of good science overruled her concerns about the costs.

Yes, perverse incentives exist, but their role as a barrier to reform should be assessed, not assumed. Ethnography’s open-ended methods, comparative frameworks and holistic explanations can capture overlooked variables and open up new avenues for action. Instead of solely blaming foot-dragging senior scientists, bring in the ethnographers.