

A crucial turning point for cognitive neuroscience

Researchers shouldn't fear papers that test and find flaws in methods. Such work contributes to better experimental designs and better science.

In 2008, Craig Bennett put a dead salmon in an magnetic resonance imaging (MRI) scanner. Bennett, a postgraduate psychology student at the University of California, Santa Barbara, then studied how the fish's brain lit up in 'response' to photographs of humans in different emotional states¹.

That this experiment discerned any brain activity at all – it was intended purely as an exercise to calibrate the scanner – served as an early warning sign that care should be taken in interpreting the statistical significance of findings from brain-imaging experiments. Fast forward to today, and some think the field of cognitive neuroscience has a full-blown reproducibility problem. Conversely, others think that the salmon study, along with subsequent work identifying methodological weaknesses, has moved the field forwards, inspiring researchers to make better decisions about experimental design and data interpretation.

In March, *Nature* published a paper² by Scott Marek at Washington University School of Medicine in St. Louis, Missouri, and his colleagues that investigated the reproducibility of brain-wide association studies. Such studies use neuroimaging techniques to explore how variations in brain structure or function affect behaviour, cognition or mental health. Marek *et al.* found that sample sizes in the thousands are needed to reliably characterize such relationships, although the authors note that they did not investigate all possible techniques or populations. The paper prompted some soul-searching that will hopefully move the field towards more robust work.

Predictive puzzles

This week, Abigail Greene at Yale University School of Medicine in New Haven, Connecticut, and her colleagues tackle the reliability of predictive modelling in cognitive neuroscience³. The method, which is used widely in the biological sciences, uses existing data sets to forecast future outcomes. It has been applied to cognitive neuroscience in an effort to determine the relationship between patterns of brain activity and various cognitive and behavioural traits. Unlike brain-wide association studies, predictive-modelling studies can be reliable with smaller sample sizes.

Greene and her co-workers systematically characterized the cases for which predictive models fail to generate accurate predictions in cognitive neuroscience, and show

that this failure is not random. Rather, it tends to occur for certain groups of people regardless of the data set – groups that aren't average.

This might be interpreted as showing that, in cognitive neuroscience, predictive models lack methodological robustness, fuelling wider concerns about the field. Some researchers have told *Nature* that, since the publication of Marek and colleagues' work, reviewers of papers and grants have had a more negative view of neuroimaging studies with small sample sizes – even if they are not brain-wide association studies. The implication is that grants need to get larger, involving consortia that can collect data from thousands, which could crowd out small research groups and researchers in low-resource settings.

Others fear that the findings will contribute to a perception among scientists outside the field that cognitive neuroscience is statistically underpowered and based on models that systematically fail. However, these studies provide the opportunity for significant growth in the field, as they have done in others.

Around 20 years ago, the genetics community needed to confront the reality that studies looking to determine the genetic basis of traits using candidate-gene approaches were not producing results that said meaningful things about genes and diseases. Genetics was much more complex than they had originally realized and, among other things, needed greater statistical firepower.

Researchers turned to genome-wide association studies, which scan the genomes of many people in an effort to determine whether and how variations are associated with particular diseases, such as heart disease or cancer. One of the earliest such studies, of 96 people with age-related macular degeneration – a major cause of blindness in older people – and 50 control participants, revealed more about the hereditary nature of the condition⁴. Studies involving much larger numbers of people soon followed, and researchers have since confirmed that larger sample sizes are better for reproducibility⁵. As a result, genetics has been transformed. It is both more robust and more collaborative, with statisticians working alongside life scientists.

The field of cognitive neuroscience has been experiencing a growth spurt similar to the one genetics went through two decades ago. Growth requires a lot of energy and can be painful, but it is an integral part of life and evolution. The findings of Greene *et al.* and Marek *et al.* should not be seen as a criticism of the field or its methods, nor be interpreted as evidence of a reproducibility crisis. By presenting clear analyses to guide researchers in choosing their experimental designs and interpreting their results when using two important methods, they provide the sort of self-reflection necessary to move cognitive neuroscience to the next level. For a discipline to progress, we must not only appreciate its strengths, but also understand its weaknesses.

“
For a discipline to progress, we must not only appreciate its strengths, but also understand its weaknesses.”

1. Bennett, C. M., Miller, M. B. & Wolford, G. L. *NeuroImage* **47**, S125 (2009).
2. Marek, S. *et al.* *Nature* **603**, 654–660 (2022).
3. Greene, A. S. *et al.* *Nature* <https://doi.org/10.1038/s41586-022-05118-w> (2022).
4. Klein, R. J. *et al.* *Science* **308**, 385–389 (2005).
5. Duncan, L. E., Ostacher, M. & Ballon, J. *Neuropsychopharmacology* **44**, 1518–1523 (2019).