UNIVERSITY OF BATH

Reboot undergraduate courses for reproducibility

Collaboration across institutes can train students in open, team science, which better prepares them for challenges to come, says **Katherine Button**.

Three years ago, as I prepared to start as a lecturer in the University of Bath's psychology department, I reflected on my own undergraduate training. What should I emulate? What would I like to improve? The 'reproducibility crisis' was in full swing. Many of the standard research practices I had been taught were now shown to be flawed, from *P*-value hacking to 'HARKing' — hypothesizing after the results are known — and an over-reliance on underpowered studies (that is, drawing oversized conclusions from undersized samples).

It struck me that the research dissertation students do in their final year is almost a bootcamp for instilling these bad habits. Vast numbers of projects, limited time and resources, small sample sizes, the potential for undisclosed analytic flexibility (*P*-hacking) and a premium on novelty: together, a recipe for irreproducible results.

Most undergraduate dissertations turn into exercises tallying the limitations of the research design — frustrating for both student and supervisor. However, each year a few students get lucky and publish, securing a huge CV advantage. I wondered what lesson this was teaching. Were we embedding a culture that rewards chance results over robust methods?

In an effort to disrupt this culture, I set up the GW4 Undergraduate Psychology Consortium with colleagues at the universities of Bath, Bristol, Cardiff and Exeter. We wanted to embed rigorous research practices into undergraduate education, incorporating procedures such as pre-registration of study protocols, designing studies with sufficient statistical power and transparent reporting of methods and results.

WERE WE **EMBEDDING** A CULTURE THAT REWARDS **CHANCE RESULTS** OVER ROBUST METHODS?

The difficulty was working out how. Rigorous research methods often take more time and resources than a student project allows. Our solution was collaboration. By working together, students could pool their efforts in data collection to reach samples sizes sufficient for meaningful analyses.

The Consortium is now entering its third year. We are still evolving, but we have settled into a productive routine. It works best if a PhD student or postdoc develops the primary research question for the undergraduates to tackle, drafts a 'bare-bones' study protocol and manages the study. Over the UK summer break, this protocol is circulated to undergraduate students (usually from two to five students at each institution), and each of them plans a secondary research question and suitable method.

At the start of the undergraduates' final year (in the first week of October), we hold the first consortium meeting, where students pitch their secondary questions and decide which will make it into the study. For example, if the main study question is on the effect of impulse-control training on reducing unhealthy food choices, an undergraduate might propose investigating whether effects are moderated by personality traits such as impulsiveness. The student will then propose a measure for assessing that trait, and propose an analysis to test their hypothesis. This way, each student has some design input, but the sample size and research integrity of the main project is retained. In addition, each student can focus on a slightly different question and so meet requirements for individual assessment. The study protocol is publicly preregistered (in our case, at the Open Science Framework at https://osf.io), and data collection runs for four months, from November to March.

In April, students present their findings to the group and collectively discuss the main study results. They reach consensus on conclusions and write up results for wider dissemination.

There are costs. Consortium studies take more time to set up and more effort to coordinate than does the standard student project. But these costs are a small price to pay for giving students the opportunity to network with peers and with researchers at other institutions, exposure to better practices and the feeling of being a valued part of a team. We academics benefit from aligning our teaching with our practice.

It is an example of how, with a bit of creative thinking, we can overcome some of the pitfalls of the current model when it comes to training the next generation to do quantitative experimental research. A handful of publications are in the works.

Both the open-science movement and the growth in online platforms for behavioural tasks and questionnaires have made it easier for psychologists to work across institutions. Using these, we can be confident we are running the same experimental procedures across sites.

Clearly, this approach is not appropriate for all types of research. It might be harder for wet-lab studies, say, in which consumables are expensive, and the idiosyncratic set up of labs makes it more challenging to standardize operating procedures. Yet working collaboratively might be even more beneficial when establishing generalizability or harmonizing methods are more difficult, especially given that students who enter graduate school can sometimes spend years trying to reproduce published work before building upon it.

Early training in collaboration might also bring comfort and creativity with regards to similar approaches later in students' research careers. Although real-world research is increasingly collaborative, it lacks conventions on how to adequately recognize and reward individuals' research input. Perhaps there are wider lessons to take from how we've designed our approach to align rigorous consortium research methods with university requirements for individual assessment. ■

Katherine Button *is a lecturer in the department of psychology at the University of Bath, UK.*
e-mail: k.s.button@bath.ac.uk