ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/128157/

This is the author’s version of a work that was submitted to / accepted for publication.

Citation for final published version:

Munafò, Marcus R., Chambers, Christopher D. , Collins, Alexandra M., Fortunato, Laura and Macleod, Malcolm R. 2020. Research culture and reproducibility. Trends in Cognitive Sciences 24 (2) , pp. 91-93. 10.1016/j.tics.2019.12.002

Publishers page: http://dx.doi.org/10.1016/j.tics.2019.12.002

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Research Culture and Reproducibility

Marcus Munafò [1], Christopher D. Chambers [2], Alexandra M. Collins [3], Laura Fortunato [4], Malcolm R. Macleod [5]

1. MRC Integrative Epidemiology Unit, School of Psychological Science, University of Bristol, Bristol, United Kingdom.
2. CUBRIC, School of Psychology, Cardiff University, Cardiff, United Kingdom.
3. Centre for Environmental Policy, Imperial College London, United Kingdom
4. Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, United Kingdom; Santa Fe Institute, Santa Fe, NM, United States of America
5. Centre for Clinical Brain Sciences, University of Edinburgh, Scotland, UK

Correspondence: marcus.munafo@bristol.ac.uk (Marcus Munafò)

Keywords: Replication Crisis; UK Reproducibility Network

There is ongoing debate regarding the robustness and credibility of published scientific research. We argue that these issues stem from two broad causal mechanisms: the cognitive biases of researchers, and the incentive structures within which researchers operate. The UK Reproducibility Network is working with researchers, institutions, funders, publishers and other stakeholders to address these issues.

There is ongoing debate regarding the extent to which research claims are robust and credible. Although this debate is not new – Charles Babbage wrote “Reflections On the Decline of Science in England” in 1830 [1] – the recent discussions can perhaps be traced to a seminal article by John Ioannidis, “Why Most Published Research Findings Are False”, published in 2005 [2]. Ultimately the focus turns on the question of how we, as researchers, can do better, in terms of the quality and robustness – and ultimately the usefulness - of our research outputs.

Modern research-intensive universities present a paradox. On the one hand, they are dynamic, vibrant institutions where researchers use cutting-edge methods to advance knowledge. On the other, their traditions, structures, and ways of working remain rooted in the 19th Century model of the independent scientist. A growing realization of this, and the impact it might have on the performance of research-intensive institutions, has led to growing interest in examining and understanding research culture.

In our view, issues of research quality stem from two main causes – scientists are subject to the same array of cognitive biases as anyone else [3]; and our ways of working and incentive structures within which we work have incrementally become so distorted that they are now harmful to the research endeavor. What is most difficult about tackling these two issues is that they are essentially invisible – cognitive biases are often unconscious, and culture is both pervasive and difficult to observe.

Insert Figure 1 about here.

The vast majority of scientists choose their career because they are passionate about their subject and excited by the possibility of advancing human knowledge. However, this passion can serve as a double-edged sword. When we are personally invested in our own research, then our ability to objectively analyze data may be negatively affected. We may see patterns in noise, suffer from confirmation bias, and so on. We have argued that open research practices – protocol preregistration, data and material sharing, the use of preprints and so on – can protect against these kinds of cognitive biases [4]. Promoting transparency in methods and data sharing should encourage greater self- and peer-appraisal of research methods. Although the conventional journal article format, with restrictions on word count and display items may not encourage this, exciting innovations are emerging that offer new approaches to scientific communication – pre-print servers, post publication peer review (e.g., F1000), the "Registered Reports" article format, and data repositories. Given these innovations, there is really no reason to provide only a partial account of one's research.

Open research also highlights the extent to which our current scientific culture relies heavily on trust. While this may have been appropriate in the 19th Century era of the independent scientist (although even that is debatable), it does not provide a strong basis for robust science in the highly-charged and competitive environment of modern science. At present, it is difficult for research consumers to know whether what is reported in an article is a complete and honest account of what was actually done and found.

A striking illustration of this comes from a 2011 paper, in which psychologists at the University of Pennsylvania showed that participants randomized to listen to "When I'm 64" by the Beatles became younger compared to those randomized to listen to a different track [5]. Not that they *felt* younger; they *became* younger. This finding was obviously false, but it was arrived at (with a significance level < 0.05) through extensive (but not untypical) flexibility in the design, conduct and analysis of the study. What Simmons and colleagues showed was that their results could be presented in two different ways – either a full account of all elements of the study design, which made it clear that the result had been arrived at through a process of over-enthusiastic analysis; or a partial, redacted account that was intended to tell the best, the least complicated, "story". Although few findings are as obviously false as this one, the point is that we usually have no way of knowing whether we are viewing a full account, or a story. We simply have to *trust* that the authors have reported their study fully and transparently.

This desire for narrative is reflected in something that many early career researchers are told – that their data need to "tell a story". Of course, it's clear what this metaphor is meant to convey – we should write in a clear and compelling way. But the focus on narrative has come to dominate to such an extent that perhaps the story matters more than the truth. As scientists, we are rarely incentivized by the system for being *right* – we are rewarded for papers, grants and so on, but not (directly) for getting the right answer. And our success in writing papers and winning grants often reflects our storytelling rather than our science.

Other metaphors that are common in science are similarly revealing, if we reflect on them. For example, we are told we should be doing "groundbreaking" research. But builders break ground when they start to construct something. As Ottoline Leyser says: "Ground-breaking is what you do when you start a building. You go into a field and you dig a hole in the ground. If you're only rewarded for ground-breaking research, there's going to be a lot of fields with a small hole in, and no buildings." [6]. Certain journals prioritize groundbreaking findings, and a publication in those can have a dramatic impact on career trajectory.

But is this the fault of the journals? There is a place for high-risk, high-return findings – those which may well be wrong but which if right would turn out to be transformative (which essentially is what groundbreaking research is). It is our *institutions* – their hiring and promotion practices – and to an extent we *ourselves* – the community of scientists – that fetishize publication in certain journals. By disproportionately lauding and rewarding high-risk, high-return activity, we risk incentivizing science in a manner similar to the way in which the banking system was incentivized prior to 2008 – the focus on high return investment vehicles that looked reliable and robust but in fact were built on sand. And that did not end well.

This returns us to the question of culture. Institutions and funders may, for example, sign DORA and exhort their promotion committee and funding panels to not consider Journal Impact Factor or similar metrics. But the final decisions are made by academics themselves who may, implicitly or explicitly, still use the flawed heuristics for quality. We often hear that Journal X or Funder Y rejected our article or grant. But (in most cases) it is not the journal or the funder that rejects us – it is *us*, as a community – the reviewers, editors, panel members and so on. We create our culture, invisible though it may be, and we therefore have it collectively within ourselves to change our culture for the better. And our institutions in particular, as repositories of this culture, can be critical in fostering change.

To help promote such change we established the UK Reproducibility Network (www.ukrn.org), with the aim of placing the UK at the leading edge, globally, for conducting and promoting rigorous research (Box 1). The network comprises grass-roots, peer-led networks of researchers at individual institutions, linked to a growing group of stakeholders (funders, publishers, learned societies, professional bodies, and other organizations that form part of the research ecosystem), and to institutions which have committed to efforts to improve the reproducibility of their research. We work to provide coordination within and across these levels, and a voice to researchers at all levels in ongoing discussions of reproducibility and culture.

Fundamentally, we need to better align our research culture with the demands of 21st Century research. We need to move away from a model that relies on *trust* in individual researchers towards one where the system is inherently *trustworthy*. This will require a focus on realigning incentives, so that what is good for scientists' careers is good for scientists, as well as recognition that excellence in research is not just generated by individuals, but by teams, departments, institutions and international collaborations. These teams require a diverse range of skills, each of which is critical to the success of the wider effort. And they require us to focus less on individuals and more on teams, and the systems and processes they work within.

In the UK there are mechanisms that could, in principle, be used to drive this change. The Research Excellence Framework (https://www.ref.ac.uk), for example, evaluates institutions on outputs, impact and environment. But the environment – perhaps the most critical element in terms of fostering a system that is inherently trustworthy – is given the least weighting. Placing greater emphasis on this would encourage institutions to focus more on creating a positive research environment. Ultimately, if we can optimize our systems and processes – our culture – then the work we produce will be more likely to be inherently trustworthy because of how it has been produced, rather than relying on us to merely trust the originators of the research.

UKRN is therefore working with researchers, institutions and stakeholders to foster coordinated culture change across the sector. For example, we can incentivize specific behaviours that we want to promote – open research practices, or success in fostering the development of early career researchers, say – by including these in promotion criteria. Our focus is on research quality, but these issues intersect with other timely issues such as diversity and harassment. Our approach is inherently collaborative rather than competitive – a rising tide lifts all boats. And to succeed it needs engagement at all levels, and for a plurality of views to be shared. We encourage UK researchers to join us, and researchers in other countries, in sharing ideas and ongoing initiatives for tackling these important and difficult issues.

These views represent those of the authors, and not necessarily the wider UKRN community.

References

1. Babbage, C. (1830). Reflections on the decline of science in England, and its causes. B. Fellowes, London.

2. Ioannidis, J.P.A. (2005). Why most published research findings are false. PLOS Medicine, 2(8), e124.

3. Bishop, D.V.M. (2019). The psychology of experimental psychologists: Overcoming cognitive constraints to improve research: The 47th Sir Frederic Bartlett Lecture. Quarterly Journal of Experimental Psychology.

4. Munafò, M.R. et a. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1, 0021.

5. Simmons, J.P. et al. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science, 22(11), 1359-1366.

6. Eastwood, B. (Producer) (2017). Ottoline Leyser: The Life Scientific. London: BBC Radio 4.

Box 1: The UK Reproducibility Network.

The UK Reproducibility Network (UKRN) is supporting a number of initiatives at various levels across the UK research system, with the goal of ensuring UK retains its place as a centre for world-leading research, by investigating the factors that contribute to robust research, promoting training activities and disseminating best practice, and working with stakeholders to ensure coordination of efforts across the sector.

Registered Reports (https://cos.io/rr/) are a form of empirical journal article in which methods and proposed analyses are pre-registered and peer-reviewed prior to research being conducted.

Accountable Replications Policy (https://royalsocietypublishing.org/rsos/replication-studies) is an initiative that commits a journal to publishing any methodologically sound replication of any previous empirical work published within the journal, regardless of the results.

Open Research Working Groups (https://osf.io/vgt3x/) are set up by researchers at their institution, and seek to make the processes and products of research as transparent, accessible and reproducible as possible.

ReproducibiliTea (https://osf.io/3ed8x/) is an initiative developed by three early-career researchers – Amy Orben, Sam Parsons and Sophia Crüwell – that aims is to build a community of researchers interested in open and reproducible research.

Octopus (https://octopus-hypothesis.netlify.com) is a novel publishing concept, developed by Alexandra Freeman, that allows hypotheses, method, results and analyses to be published as they are produced.

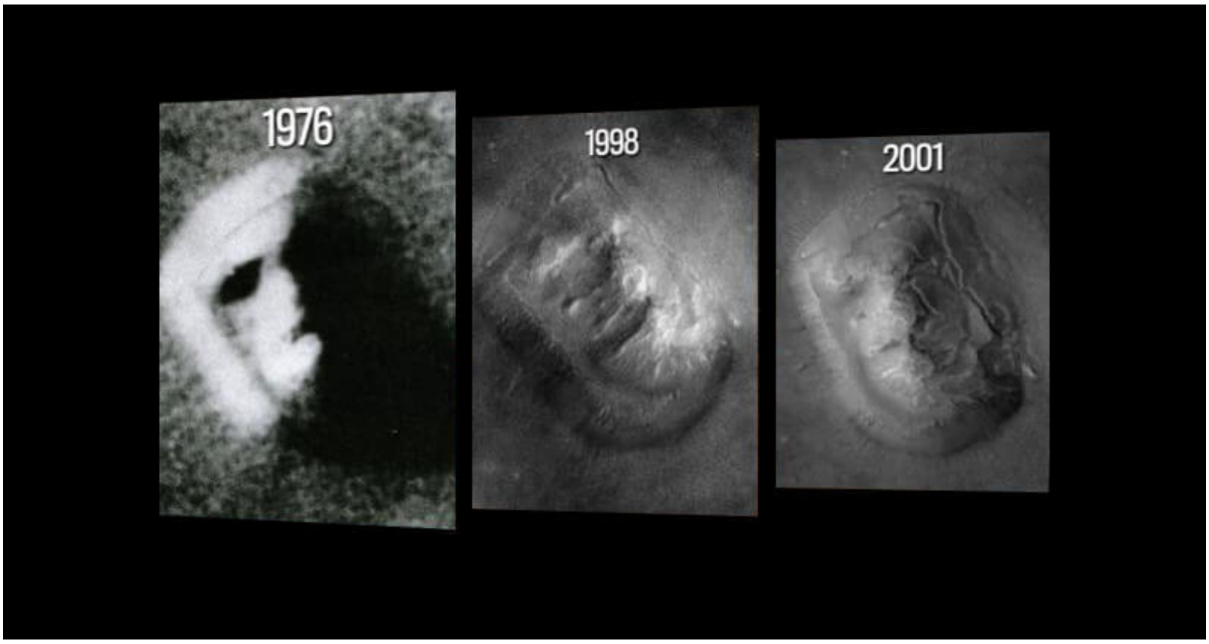
Framework for Open and Reproducible Research Training (https://forrt.netlify.com/) aims to support the teaching of open and reproducible research practices with teaching resources and a framework capturing the aspects of open and reproducible research taught within a given course.

Consortium-Based Student Projects (https://osf.io/74ur2/) is a collaborative format for research dissertation projects, where the aim is to pool resources and effort across universities so that students can participate in high-quality, well-powered research, with an open science ethos.

Laboratory Efficiency Assessment Framework (https://www.ucl.ac.uk/greenucl/resources/labs/leaf-laboratory-efficiency-assessment-framework) is an initiative to improve the efficiency and sustainability of laboratories, with a focus on reproducible research to reduce research waste.

Open Research Primers (https://www.bristol.ac.uk/psychology/research/ukrn/about/resources/) are a set of crowdsourced primers on five topics: preprints, preregistration/Registered Reports, open code/software, data sharing, and open access publishing. Each primer is written by a cross-disciplinary team.

Figure 1. A Face on Mars?



Humans have a tendency to see patterns in noise (and in particular, faces in inanimate objects – pareidolia). The first photograph taken of this geological feature on the surface of Mars looked like a face. Unfortunately, that finding failed to replicate. Scientists working with complex, noisy biological data, who are motivated and incentivized to *find* something, may be led astray by our natural, human cognitive biases. Source: Wikipedia.