

VINTAGE

Contents

Preface

PART I: OUGHT AND IS

- 1 How Science Works
- 2 The Replication Crisis

PART II: FAULTS AND FLAWS

- 3 Fraud
- 4 Bias
- 5 Negligence
- 6 Hype

PART III: CAUSES AND CURES

- 7 Perverse Incentives
- 8 Fixing Science

Epilogue

Appendix: How to Read a Scientific Paper

Acknowledgements

Notes

Index

ABOUT THE AUTHOR

Dr Stuart Ritchie is a Lecturer at the Institute of Psychiatry, Psychology and Neuroscience at King's College London and winner of the 2015 'Rising Star' award from the Association for Psychological Science. He has written for *The Times*, *Spectator*, *Washington Post*, *Wired*, *Literary Review* and *Aeon*, and has appeared on BBC Radio 4 programmes *The Infinite Monkey Cage*, *More or Less* and *Bringing Up Britain*. His Twitter account is @StuartJRitchie.

For Katharine

The author gratefully acknowledges permission to reproduce images from the following sources:

1 Sidney Harris, ScienceCartoonsPlus.com;

2 the Getty's Open Content Program;

3 Cambridge University Press;

4 Joel Pett.

The figures on 5 and 6 are reproduced under CC-BY and CC-0 Creative Commons licences respectively.

Now *that* is scientific fact. There's no real *evidence* for it, but it is scientific fact.

Brass Eye

Preface

It is the peculiar and perpetual error of the human understanding to be more moved and excited by affirmatives than by negatives.

Francis Bacon, *Novum Organum* (1620)

January 31, 2011 was the day the world found out that undergraduate students have psychic powers.

A new scientific paper had hit the headlines: a set of laboratory experiments on over 1,000 people had found evidence for psychic precognition – the ability to see into the future using extrasensory perception. This wasn't the work of some unknown crackpot: the paper was written by a top psychology professor, Daryl Bem, from the Ivy League's Cornell University. And it didn't appear in an obscure outlet – it was published in one of the most highly regarded, mainstream, peer-reviewed psychology journals.¹ Science seemed to have given its official approval to a phenomenon that hitherto had been considered completely impossible.

At the time, I was a PhD student, studying psychology at the University of Edinburgh. I dutifully read Bem's paper. Here's how one of the experiments worked. Undergraduate students looked at a computer screen, where two images of curtains would appear. They were told that there was another picture behind one of the curtains, and that they had to click whichever they thought it was. Since they had no other information, they could only guess. After they'd chosen, the curtain disappeared and they saw whether they'd been correct. This was repeated thirty-six times, then the experiment was over. The results were quietly stunning. When a picture of some neutral, boring object like a chair was behind one of the curtains, the outcome was almost perfectly random: the students chose correctly 49.8 per cent of the time, essentially fifty-fifty. However – and here's where it gets strange – when one of the pictures was pornographic, the students tended to choose it slightly more often than chance: 53.1 per cent of the time, to be exact. This met the threshold for 'statistical significance'. In his paper, Bem suggested that some unconscious, evolved, psychic sexual desire had ever-so-slightly nudged the students towards the erotic picture even before it had appeared on screen.²

Some of Bem's other experiments were less explicit, but no less puzzling. In one of them, a list of forty unrelated words appeared on the screen, one at a time. Afterwards came a surprise memory test, where the students had to type in as many of the words as they could remember. At that point, the computer randomly selected twenty of the words and showed them to the students again. Then the experiment ended. Bem reported that, during the memory test, the students were more likely to remember the twenty words *they were about to see again*, even though they couldn't have known – except by psychic intuition – which ones they were going to be shown. This would be a

bit like studying for an exam, sitting the exam, then studying again afterwards, and that post-exam study somehow winding its way back in time to improve your grade. Unless the laws of physics had suddenly been repealed, time is supposed to run in only one direction; causes are supposed to come before, not after, their effects. But with the publication of Bem's paper, these bizarre results were now a part of scientific literature.

Crucially, Bem's experiments were extremely simple, requiring nothing more complicated than a desktop computer. If Bem was right, any researcher could produce evidence for the paranormal just by following his experimental instructions – even a PhD student with next to no resources. That is what I was, so that is exactly what I did. I got in touch with two other psychologists who were also sceptical of the results, Richard Wiseman of the University of Hertfordshire and Chris French of Goldsmiths, University of London. We agreed to re-run Bem's word-list experiment three times, once at each of our respective universities. After a few weeks of recruiting participants, waiting for them to complete the memory test and then dealing with their looks of bewilderment as we explained afterwards what we'd been looking for, we had the results. They showed ... nothing. Our undergraduates weren't psychic: there was no difference in their recall of the words presented after the test. Perhaps the laws of physics were safe after all.

We duly wrote up our results and sent the resulting paper off to the same scientific journal that had published Bem's study, the *Journal of Personality and Social Psychology*. Almost immediately the door was slammed in our faces. The editor rejected the paper within a few days, explaining to us that they had a policy of never publishing studies that repeated a previous experiment, whether or not those studies found the same results as the original.³

Were we wrong to feel aggrieved? The journal had published a paper that had made some extremely bold claims – claims that, if true, weren't just interesting to psychologists, but would completely revolutionise science. The results had made their way into the public domain and received significant publicity in the popular media, including an appearance by Bem on the late-night talk show *The Colbert Report* where the host coined the memorable phrase 'time-travelling porn'.⁴ Yet the editors wouldn't even consider publishing a replication study that called the findings into question.⁵

Meanwhile, another case was unfolding that also raised alarming questions about the current state of scientific practice. *Science*, widely considered one of the world's most prestigious scientific journals (second only to *Nature*), had published a paper by Diederik Stapel, a social psychologist at Tilburg University in the Netherlands. The paper, entitled 'Coping with Chaos', described several studies performed in the lab and on the street, finding that people showed more prejudice – and endorsed more racial stereotypes – when in a messier or dirtier environment.⁶ This, and some of Stapel's dozens of other papers, hit the headlines across the world. 'Chaos Promotes Stereotyping', wrote *Nature's* news service; 'Where There's Rubbish There's Racism', alliterated the *Sydney Morning Herald*.⁷ The results exemplified a type of social psychology research that produced easy-to-grasp findings with, as Stapel himself wrote, 'clear policy implications': in this case, to 'diagnose environmental disorder early and intervene immediately'.⁸

The problem was that none of it was real. Some of Stapel's colleagues became suspicious after they noticed the results of his experiments were a little *too* perfect.

Not only that, but whereas senior academics are normally extremely busy and rely on their students to do such menial tasks as collecting data, Stapel had apparently gone out and collected all the data himself. After the colleagues brought these concerns to the university in September 2011, Stapel was suspended from his professorship. Multiple investigations followed.⁹

In a confessional autobiography he wrote subsequently, Stapel admitted that instead of collecting the data for his studies, he would sit alone in his office or at his kitchen table late into the night, typing the numbers he required for his imaginary results into a spreadsheet, making them all up from scratch. 'I did some things that were terrible, maybe even disgusting,' he wrote. 'I faked research data and invented studies that had never happened. I worked alone, knowing exactly what I was doing ... I didn't feel anything: no disgust, no shame, no regrets.'¹⁰ His scientific fraud was surprisingly elaborate. 'I invented entire schools where I'd done my research, teachers with whom I'd discussed the experiments, lectures that I'd given, social-studies lessons that I'd contributed to, gifts that I'd handed out as thanks for people's participation.'¹¹

Stapel described printing off the blank worksheets he'd ostensibly be giving to his participants, showing them to his colleagues and students, announcing he was heading off to run the study ... then dumping the sheets into the recycling when nobody was looking. It couldn't last. The findings of the investigations were clear; he was fired not long after his suspension. Since then, no fewer than fifty-eight of his studies have been retracted – struck off the scientific record – due to their fake data.

The Bem and Stapel cases – where esteemed professors published seemingly impossible results (in Bem's case) and outright fraudulent ones (in Stapel's) – sent a jolt through psychology research, and through science more generally. How could prestigious scientific journals have allowed their publication? How many other studies had been published that couldn't be trusted? It turned out that these cases were perfect examples of much wider problems with the way we do science.

In both cases, the central issue had to do with *replication*. For a scientific finding to be worth taking seriously, it can't be something that occurred because of random chance, or a glitch in the equipment, or because the scientist was cheating or dissembling. It has to have really happened. And if it did, then in principle I should be able to go out and find broadly the same results as yours. In many ways, that's the essence of science, and something that sets it apart from other ways of knowing about the world: if it won't replicate, then it's hard to describe what you've done as scientific at all.

What was concerning, then, wasn't so much that Bem's experiments were unreliable or that Stapel's were a figment of his imagination: some missteps and spurious results will always be with us (and so, alas, will fraudsters).¹² What was truly problematic was how the scientific community had handled both situations. Our attempted replication of Bem's experiment was unceremoniously rejected from the journal that published the original; in the case of Stapel, nobody had ever even *tried* to replicate his findings. In other words, the community had demonstrated that it was content to take the dramatic claims in these studies at face value, without checking how durable the results really were. And if there are no double-checks on the replicability of results, how do we know they aren't just flukes or fakes?

Perhaps Bem himself best summed up many scientists' attitudes to replication, in an interview some years after his infamous study. 'I'm all for rigor,' he said, 'but I don't have the patience for it ... If you looked at all my past experiments, they were always

rhetorical devices. I gathered data to show how my point would be made. I used data as a point of persuasion, and I never really worried about, “Will this replicate or will this not?”¹³

Worrying about whether results will replicate or not isn't optional. It's the basic spirit of science; a spirit that's supposed to be made manifest in the system of peer review and journal publication, which acts as a bulwark against false findings, mistaken experiments and dodgy data. As this book will show, though, that system is badly broken. Important knowledge, discovered by scientists but not deemed interesting enough to publish, is being altered or hidden, distorting the scientific record and damaging our medicine, technology, educational interventions and government policies. Huge resources, poured into science in the expectation of a useful return, are being wasted on research that's utterly uninformative. Entirely avoidable errors and slip-ups routinely make it past the Maginot Line of peer review. Books, media reports and our heads are being filled with 'facts' that are either incorrect, exaggerated, or drastically misleading. And in the very worst cases, particularly where medical science is concerned, people are dying.

Other books feature scientists taking the fight to a rogue's gallery of pseudoscientists: creationists, homeopaths, flat-Earthers, astrologers, and their ilk, who misunderstand and abuse science – usually unwittingly, sometimes maliciously, and always irresponsibly.¹⁴ This book is different. It reveals a deep corruption within science itself: a corruption that affects the very culture in which research is practised and published. Science, the discipline in which we should find the harshest scepticism, the most pin-sharp rationality and the hardest-headed empiricism, has become home to a dizzying array of incompetence, delusion, lies and self-deception. In the process, the central purpose of science – to find our way ever closer to the truth – is being undermined.

The book begins by showing, in Part I, that doing science involves much more than just running experiments or testing hypotheses. Science is inherently a *social* thing, where you have to convince other people – other scientists – of what you've found. And since science is also a *human* thing, we know that any scientists will be prone to human characteristics, such as irrationality, biases, lapses in attention, in-group favouritism and outright cheating to get what they want. To enable scientists to convince one another while trying to transcend the inherent limitations of human nature, science has evolved a system of checks and balances that – in theory – sorts the scientific wheat from the chaff. This process of scrutiny and validation, which leads to the supposed gold-standard of publication in a peer-reviewed scientific journal, is described in Chapter 1. But Chapter 2 shows that the process must have gone terribly wrong: there are numerous published findings across many different areas of science that can't be replicated and whose truth is very much in doubt.

Then, in Part II, we'll ask why. We'll discover that our publication system, far from neutralising or overriding all the human problems, allows them to leave their mark on the scientific record – and does so precisely because it *believes* itself to be objective and unbiased. A peculiar complacency, a strange arrogance, has taken hold, where the mere existence of the peer-review system seems to have stopped us from recognising its flaws. Peer-reviewed papers are supposedly as near as one can get to an objective factual account of how the world works. But in our tour through many dozens of those papers, we'll discover that peer review can't be relied upon to ensure scientists are

honest (Chapter 3), detached (Chapter 4), scrupulous (Chapter 5), or sober (Chapter 6) about their results.

Part III digs deeper into scientific practice. Chapter 7 shows that it's not just that the system fails to deal with all the kinds of malpractice we've discussed. In fact, the way academic research is currently set up *incentivises* these problems, encouraging researchers to obsess about prestige, fame, funding and reputation at the expense of rigorous, reliable results. Finally, after we've diagnosed the problem, Chapter 8 describes a set of often-radical reforms to scientific practice that could help reorient it towards its original purpose: discovering facts about the world.

To make the case about the frailties of scientific research, throughout the book I'll draw on cautionary tales from a wide variety of scientific fields. Partly because I'm a psychologist, there'll be a preponderance of examples from that subject.¹⁵ My background isn't the only reason there's so much psychology in the book: it's also because after the Bem and Stapel affairs (among many others), psychologists have begun to engage in some intense soul-searching. More than perhaps any other field, we've begun to recognise our deep-seated flaws and to develop systematic ways to address them – ways that are beginning to be adopted across many different disciplines of science.

The first step in fixing our broken scientific system is learning to spot, and correct, the mistakes that can lead it astray. And the only way to do this is with more science. Throughout the book, I'll draw on *meta-science*: a relatively new kind of scientific research that focuses on scientific research itself. If science is the process of exposing and eliminating errors, meta-science represents that process aimed inwards.

Much can be learned from mistakes. On one of his albums, the musician Todd Rundgren has a spoken-word introduction encouraging the listener to play a little game he calls 'Sounds of the Studio'. Rundgren describes all the missteps that can be made when recording music: hums, hisses, pops on the microphone whenever someone sings a word with a 'p' in it, choppy editing, and so on. He suggests that the reader listen for these mistakes in the songs that follow, and on other records. And just as a better understanding of recording studio slip-ups can give you a new insight into how music is made, learning about how science goes wrong can tell you a lot about the process by which we arrive at our knowledge.

Discovering the serious problems with the way we do science will be disconcerting. How many intriguing results that you've read about in the news and popular science books, or seen in documentaries – discoveries you've been excited enough to share with friends, or that made you rethink how the world works – are based on weak research that can't be replicated? How many times has your doctor prescribed you a drug or other treatment that rests on flawed evidence? How many times have you changed your diet, your purchasing habits, or some other aspect of your lifestyle on the basis of a scientific study, only for the evidence to be completely overturned by a new study a few months later? How many times have politicians made laws or policies that directly impact people's lives, citing science that won't stand up to scrutiny? In each case, the answer is: a lot more than you'd like to think.

It's naïve to hope that every single scientific study will be true – that is, a report of ironclad facts that will never be overturned in future research. The world is far too messy a place for that. All we can hope for is that our scientific studies are trustworthy – that they honestly report what occurred in the research. If the much-vaunted peer-review process can't justify that trust, science loses one of its most basic and most

desirable qualities, along with its ability to do what it does best: revolutionise our world with a steady progression of new discoveries, technologies, treatments and cures.

I come to praise science, not to bury it; this book is anything but an attack on science itself, or on its methods. Rather, it is a defence of those methods, and of scientific principles more generally, against the way science is currently practised. What makes all the disasters we'll encounter so disturbing is the importance of science: by allowing it to become so tarnished, and its progress to be so badly stalled, we're in danger of ruining one of the greatest accomplishments of our species.

But the damage isn't irreparable. In principle, if not in practice, science still has the potential to be the robust and reliable system of knowledge we need it to be. As we explore the litany of scientific failures in the book, the positive thought to hold onto – the fragile scrap of hope and reassurance that emerges from the Pandora's box of fraud, bias, negligence and hype that we'll prise open in what follows – is that nearly all of these problems have been uncovered *by other scientists*. The clever meta-scientific ideas that have been proposed to combat these problems and clear up the mess that has been created have come, in substantial part, from within the scientific community. Even if it's been deeply buried in many fields, the self-critical spirit that animates genuine science remains.

And that's just as well, because as we're about to find out, there really is *quite* a mess.

PART I

UGHT AND IS

How Science Works

Such subjects of thought furnish not sufficient employment in solitude, but require the company and conversation of our fellow-creatures, to render them a proper exercise for the mind.

David Hume, 'Of Essay-Writing' (1777)

Science is a social construct.

Before that statement makes you toss the book across the room, let me explain what I mean. I don't mean it in the sense used by extreme relativists, post-modernists, anti-science crusaders, and others who suggest that there's no real world out there, that science is only one not-particularly-special way of knowing about it, or even that science is just one 'myth' among many that we could choose to believe.¹ Science has cured diseases, mapped the brain, forecasted the climate, and split the atom; it's the best method we have of figuring out how the universe works and of bending it to our will. It is, in other words, our best way of moving towards the truth. Of course, we might never fully get there – a glance at history shows how hubristic it is to claim any facts as absolute or unchanging. For ratcheting our way towards better knowledge about the world, though, the methods of science are as good as it gets.

But we can't make progress with those methods alone. It's not enough to make a solitary observation in your lab; you must also convince other scientists that you've discovered something real. This is where the social part comes in. Philosophers have long discussed how important it is for scientists to *show* their fellow researchers how they came to their conclusions. John Stuart Mill puts it this way:

In natural philosophy, there is always some other explanation possible of the same facts; some geocentric theory instead of heliocentric, some phlogiston instead of oxygen; and it has to be shown why that other theory cannot be the true one: and until this is shown, and until we know how it is shown, we do not understand the grounds of our opinion.²

And so, scientists work together in teams, travel the world to give lectures and conference speeches, debate each other in seminars, form scientific societies to share research and, perhaps most importantly, publish their results in peer-reviewed journals. These social aspects aren't just a perk of the job, nor mere camaraderie. They're the process of science in action: an ongoing march of collective scrutiny, questioning, revision, refinement and *consensus*. Although it might sound paradoxical

at first, the subjective process of science is what provides it with its unmatched degree of objectivity.³

It's in this sense, then, that science is a social construct. Any claim about the world can only be described as scientific knowledge after it's been through this communal process, which is designed to sieve out errors and faults and allow other scientists to say whether they judge a new finding to be reliable, robust and important. That each discovery has to run such a gauntlet imbues the eventual products of the scientific process – the published, peer-reviewed studies – with a great deal of power in society. This is no mere cant, rhetoric, or opinion, we say: this is *science*.

Science's social nature does come with weaknesses, however. Because scientists focus so much on trying to persuade their peers, which is the way they get those studies through peer review and onward to publication, it's all too easy for them to disregard the real object of science: getting us closer to the truth. And because scientists are human beings, the ways that they try to persuade each other aren't always fully rational or objective.⁴ If we don't take great care, our scientific process can become permeated by very human flaws.

This book is about how we haven't taken enough care of our precious scientific process. It's about how we ended up with a scientific system that doesn't just overlook our human foibles, but amplifies them. In recent years, it's become increasingly, painfully obvious that peer review is far from the guarantee of accuracy and reliability it's cracked up to be, while the system of publication that's supposed to be a crucial strength of science has become its Achilles' heel.

To understand how the scientific publication system has gone so wrong, though, we first need to know how it's supposed to work when it goes right.

Let's imagine you want to do some science. The first step is to read the scientific literature. This consists of a vast library of journals, the specialist magazines that are the main outlets for new scientific knowledge. The idea of a periodical where scientists could share their work dates back to 1665, when Henry Oldenburg of the UK's Royal Society published the first issue of, to give it its full title, *Philosophical Transactions: Giving Some Account of the Present Undertakings, Studies, and Labours of the Ingenious in Many Considerable Parts of the World*.⁵ The intention was that those ingenious scientists could send in letters describing their exploits, for the perusal of other interested readers. Before that, scientists either laboured alone in the courts of wealthy rulers or for private patrons or guilds (where their science was often seen as more akin to a parlour trick than an effort to discover the truth), published standalone books, or formed letter-writing circles with like-minded peers. Indeed, this latter kind of correspondence club is where institutions like the Royal Society originated.⁶

The initial issues of Oldenburg's journal were more like a newsletter, with descriptions of recent experiments and discoveries. For example, Volume 1, Issue 1 described the first ever observation of what was probably the Great Red Spot of Jupiter, by the natural philosopher and polymath Robert Hooke. The entire entry read:

The Ingenious Mr. Hook did, some months since, intimate to a friend of his, that he had, with an excellent twelve foot Telescope, observed, some days before, he than spoke of it, (*videl.* on the ninth of May, 1664, about 9 of the clock at night) a small Spot in the biggest of the 3 obscurer Belts of *Jupiter*, and that, observing it

from time to time, he found, that within 2 hours after, the said Spot had moved from East to West, about half the length of the Diameter of *Jupiter*.⁷

The journal still exists to this day, with the somewhat easier-to-remember title of *Philosophical Transactions of the Royal Society*.⁸ As time went on, the brief news items were replaced with longer articles containing detailed descriptions of experiments and studies. It's now part of a global ecosystem of over 30,000 journals, ranging from the very general (like the highly prestigious journals *Nature* and *Science*, which aim to publish the world's most noteworthy research from any scientific field) to the very specific (like the *American Journal of Potato Research*, which is only interested in papers about one tuberous topic in particular).⁹ Some journals, like *Philosophical Transactions*, are still run by scientific societies, but most are owned by commercial outfits such as Elsevier, Wiley and Springer Nature.¹⁰ A recent advancement is that scientific journals are all online, allowing anyone who can afford to pay the publisher's subscription fees – or have their university library do so on their behalf – to have the world's scientific knowledge at their fingertips.¹¹

After reading the journals relevant to your field, you might alight on a research question. Maybe there's a scientific theory that makes a prediction – an hypothesis – that you can test in some clever way; maybe there's a gap in our existing knowledge that you know just how to plug; maybe you've had a spark of inspiration and have come up with an experiment that tests something entirely new. Before you can do any of this, though, you'll normally need some money to fund the study: for instance, to buy new equipment or materials, to recruit participants, or to pay the salaries of the scientists you'll hire to do the legwork. Unless you happen to be, say, a pharmaceutical company that can afford to run its own laboratories, the main way to get that all-important funding is to apply for a grant. This might come from your government, a business, an endowment fund, a non-profit, a charity, or even a wealthy individual. You might apply to the National Institutes of Health or the National Science Foundation (both of which are taxpayer-funded agencies in the United States), or to a science-funding charity like the Wellcome Trust or the Bill & Melinda Gates Foundation.¹²

Funding is by no means assured, and any scientist will tell you that one of the most gruelling parts of the job is trying to get their latest research ideas funded, with failure grindingly common. This grasping for cash has important knock-on effects on the science itself, and we'll return to them later in the book. But for now, let's imagine you're successful in securing a grant. You can then get to work. Collecting the data might involve smashing particles together in an underground collider, finding fossils in the rocks of the Canadian Arctic, setting up the precise environment for bacterial growth in a petri dish, organising hundreds of people to come to a lab and fill in questionnaires, or running a complex computer model; it can take days, months, decades.

Once the data are in, you'll normally have a set of numbers that you, or a more mathematically minded colleague, can analyse using some variety of statistics (another minefield to which we'll return). Then you need to write it all up in the form of a scientific paper. The typical paper starts with an Introduction, where you summarise what's known on the topic and what your study adds. There follows a Method section, where you describe exactly what you did – in enough detail so that anyone could, in theory, run exactly the same experiment again. You'll then move on to a Results section, where you present the numbers, tables, graphs and statistical analyses that

document your findings, and you'll end with a Discussion section where you speculate wildly – er, I mean, provide thoughtful, informed consideration – about what it all means. You'll top the whole thing with an Abstract: a brief statement, usually of around 150 words, that summarises the whole study and its results. The Abstract is always available for anyone to read, even if the full paper is behind the journal's subscription paywall, so you'll want to use it to make your results sound compelling. Papers come in all lengths and sizes, and sometimes mix up the above order, but in general your paper will end up along these lines.¹³

When the paper is ready, you enter the world of scientific journals, and the competition for publication. Until recently, submitting a paper to a journal meant printing out several hard copies and mailing them to the editor, but nowadays everything is handled online – though many journals still use such archaic, buggy websites that you might as well send your paper by carrier pigeon. The journal's editor, often a senior academic, will read the paper (or, let's be honest, probably just the Abstract) and decide whether it might be worth publishing. Most journals, especially the highly prestigious ones, pride themselves on their exclusivity and thus their low acceptance rate (*Science*, for example, accepts less than 7 per cent of submissions), so the majority of papers will be bounced back to the authors at this point, in what's called a 'desk rejection'.¹⁴ This is the initial step in quality control: a sorting by the editor of the papers into those that match the theme of the journal and have potential in terms of their scientific interest or quality, and those that aren't worth a second look. For the fraction of articles that do take the editor's fancy, now comes the moment of peer review. The editor will find two or three scientists who are experts in your field of research and ask them whether they'd like to evaluate your manuscript. They'll probably decline because they're too busy, so the editor will keep going down the list of possible reviewers until a few agree. And so begins the nail-biting wait to see if your work will receive their endorsement.

Most people, including scientists, assume peer review has always been a crucial feature of scientific publication, but its history is more complicated. Although in the seventeenth century the Royal Society tended to ask some of its members whether they thought a paper was interesting enough to publish in *Philosophical Transactions*, requiring them to provide a written evaluation of each study wasn't tried until at least 1831.¹⁵ Even then, the formal peer review system we know today didn't become universal until well into the twentieth century (as you can tell from a letter Albert Einstein sent in 1936 to the editors of *Physical Review*, huffily announcing that he was withdrawing his paper from consideration at their journal because they had dared to send it to another physicist for comment).¹⁶ It took until the 1970s for all journals to adopt the modern model of sending out submissions to independent experts for peer review, giving them the gatekeeping role they have today.¹⁷

Peer reviewers are usually anonymous, which is both a blessing and a curse: a blessing because it allows them to speak their minds without concern about repercussions from the scientists whose work they're criticising (a junior scientist can be truly honest about the flaws of a big-name professor's work), but a curse because, well, it allows them to speak their minds without concern about repercussions from the scientists whose work they're criticising. The following are genuine excerpts from peer reviews:

- Some papers are a pleasure to read. This is not one of them.'
- The results are as weak as a wet noodle.'

- The manuscript makes three claims: The first we've known for years, the second for decades, the third for centuries.'
- I am afraid this manuscript may contribute not so much towards the field's advancement as much as towards its eventual demise.'
- Did you have a seizure while writing this sentence? Because I feel like I had one while reading it.'¹⁸

If the reviewers' evaluations look like this, the editor will probably reject your paper. At that point you might want to give up, or start the whole process again by submitting to a different journal, and if that fails a different one, and if that fails a different one, and so on – it's not uncommon for papers to go through half a dozen or more journals, usually of ever-lower prestige, before they get accepted for publication. If, on the other hand, the reviewers are more impressed, you might get the opportunity to revise your paper to respond to their critiques – perhaps running new analyses or new experiments, or rewriting certain sections – and submit it to the editor again. The back-and-forth revising process can go through multiple rounds, and often takes months. Eventually, if the reviewers are satisfied, the editor gives the go-ahead and the paper is published. If the journal still prints hard copies, you'll get to see your precious work in print; otherwise, you'll have to settle for the thrill of seeing it on the journal's official website. That's it. You've made your mark on the scientific literature, and you have a publication that you can add to your CV and that can be cited by other researchers. Congratulations – take the rest of the day off.

The above summary is all too brief and general, but essentially every scientific field follows that process in some form. We might ask ourselves whether, after being put through the mangle of peer review, the eventual publication still provides a faithful representation of what was done in the study. We'll get to that in later chapters. For now, we need to consider something else. What ensures that the participants in the process just described – the researcher who submits the paper, the editor at the journal, the peers who review it – all conduct themselves with the honesty and integrity that trustworthy science requires? There's no law requiring that everyone acts fairly and rationally when evaluating science, so what's needed is a shared ethos, a set of values that aligns the scientists' behaviour.¹⁹ The best-known attempt to write down these unwritten rules is that of the sociologist Robert Merton.

In 1942, Merton set out four scientific values, now known as the 'Mertonian Norms'. None of them have snappy names, but all of them are good aspirations for scientists. First, *universalism*: scientific knowledge is scientific knowledge, no matter who comes up with it – so long as their methods for finding that knowledge are sound. The race, sex, age, gender, sexuality, income, social background, nationality, popularity, or any other status of a scientist should have no bearing on how their factual claims are assessed. You also can't judge someone's research based on what a pleasant or unpleasant person they are – which should come as a relief for some of my more disagreeable colleagues. Second, and relatedly, *disinterestedness*: scientists aren't in it for the money, for political or ideological reasons, or to enhance their own ego or reputation (or the reputation of their university, country, or anything else). They're in it to advance our understanding of the universe by discovering things and making things – full stop.²⁰ As Charles Darwin once wrote, a scientist 'ought to have no wishes, no affections, – a mere heart of stone.'²¹

The next two norms remind us of the social nature of science. The third is *communality*: scientists should share knowledge with each other.²² This principle

underlies the whole idea of publishing your results in a journal for others to see – we’re all in this together; we have to know the details of other scientists’ work so that we can assess and build on it.²³ Lastly, there’s *organised scepticism*: nothing is sacred, and a scientific claim should never be accepted at face value. We should suspend judgement on any given finding until we’ve properly checked all the data and methodology. The most obvious embodiment of the norm of organised scepticism is peer review itself.

It looks good in theory: by following the four Mertonian Norms, we should end up with a scientific literature we can trust – the shoulders of giants, as in Newton’s famous phrase, on which we stand to see farther. Of course, those giants often had it wrong: just to take the two examples mentioned above by John Stuart Mill, we used to believe that the Sun orbited the Earth, and that flammable objects were full of a special element called phlogiston that was released when they burned.²⁴ But we eventually consigned these theories to the scrapheap as better data came in. Indeed, it’s a virtue for a scientist to change their mind. The biologist Richard Dawkins recounts his experience of ‘a respected elder statesman of the Zoology Department at Oxford’ who for years had:

passionately believed, and taught, that the Golgi Apparatus (a microscopic feature of the interior of cells) was not real: an artefact, an illusion. Every Monday afternoon it was the custom for the whole department to listen to a research talk by a visiting lecturer. One Monday, the visitor was an American cell biologist who presented completely convincing evidence that the Golgi Apparatus was real. At the end of the lecture, the old man strode to the front of the hall, shook the American by the hand and said – with passion – “My dear fellow, I wish to thank you. I have been wrong these fifteen years.” We clapped our hands red ... In practice, not all scientists would [say that]. But all scientists pay lip service to it as an ideal – unlike, say, politicians who would probably condemn it as flip-flopping. The memory of the incident I have described still brings a lump to my throat.²⁵

This is what people mean when they talk about science being ‘self-correcting’. Eventually, even if it takes many years or decades, older, incorrect ideas are overturned by data (or sometimes, as was rather morbidly noted by the physicist Max Planck, by all their stubborn proponents dying and leaving science to the next generation).²⁶ Again, that’s the theory. In practice, though, the publication system described earlier in this chapter sits awkwardly with the Mertonian Norms, in many ways obstructing the process of self-correction. The specifics of this contradiction – between the competition for grants and clamour for prestigious publications on the one hand, and the open, dispassionate, sceptical appraisal of science on the other – will become increasingly clear as we progress through the book.

For now, though, notice what it was that changed the mind of Dawkins’s elder statesman: ‘completely convincing evidence’. There’s little point in trying to correct and update our scientific theories with data if the data themselves aren’t convincing – or worse, aren’t even accurate. This brings us back to the idea we discussed in the Preface: for results to warrant our trust, they need to be replicable. As the philosopher of science Sir Karl Popper puts it:

Index

The page references in this index correspond to the print edition from which this ebook was created, and clicking on them will take you to the location in the ebook where the equivalent print page would begin. To find a specific word or phrase from the index, please use the search feature of your ebook reader.

ABC News, 170
abortion, 90
Abu Ghraib prison abuse (2003), 30
accidental discoveries, 212
Acta Crystallographica Section E, 182
acupuncture, 70
Afghan hounds, 56, 57
Agence France-Presse, 223
AIDS (acquired immune deficiency syndrome), 239, 244
Alchemist, The (Bega), 196, 197
Alexander, Benita, 51
Alexander, Scott, 141
algorithms, 35, 126, 203, 232, 240
allergies, 39–40
Alzheimer, Aloysius, 114
Alzheimer's Disease, 114–15, 160, 161
Amazon, 35
American Journal of Potato Research, 16
Amgen, 37
amygdala, 140
amyloid cascade hypothesis, 114–15
anaesthesia
 awareness, 39, 41
 Fujii affair (2012), 67–8, 129
 outcome switching, 112
Anaesthesia & Analgesia, 68
animal studies, 133–4, 137, 143, 149–50
antidepressants, 95, 184, 199
antipsychotics, 107, 139
archaeology, 93, 95
Arnold, Frances, 67
arsenic, 146–8, 170
artificial tracheas, 48–54
asthma, 139

austerity, 123–5
 Australia, 70, 124, 178, 242
 Austria, 124
 autism, 76–9, 161–3
 aviation, 132–3

Babbage, Charles, 241
 Bacon, Francis, 1
 bacteria, 145–8, 161–4
 Bargh, John, 154–5
 Bayer, 37
 Bayes, Thomas, 207
 Bayesian statistics, 207–8
 BDNF gene, 140
Before You Know It (Bargh), 154–5
 Bega, Cornelis, 196, 197
 Begley, Sharon, 114–15
 Belgium, 124
 Bell Labs, 71, 72
 Bem, Daryl, 1–3, 5, 6, 8, 27, 31, 42, 176, 205, 240
 benzodiazepines, 139
 bias, 81–121, 200, 251

- blinding and, 134
- conflict of interest, 113, 115–16, 251
- De Vries' study (2018), 199, 201
- funding and, 112–13
- groupthink and, 114
- meaning well bias, 113
- Morton's skull studies, 81–3, 119–21, 253
- p*-hacking, 99–112, 113
- politics and, 115–17, 119
- publication bias, 86–112, 116, 165, 196, 199–200, 204–6
- randomisation and, 110, 129
- sexism and, 117–18

Bik, Elisabeth, 60–61, 69, 203
 Bill & Melinda Gates Foundation, 17, 219
Biomaterials, 49
 biology

- amyloid cascade hypothesis, 114–15
- Bik's fake images study (2016), 60–61, 69
- Boldt affair (2010), 76
- cell lines, 131–3, 143
- China, misconduct in, 70
- Hwang affair (2005–6), 55–7, 60, 74, 202
- Macchiarini affair (2015–16), 48–54, 69, 80, 202
- meta-scientific research, 112
- microbiome studies, 161–4
- Morton's skull studies, 81–3, 119–21, 253

- Obokata affair (2014), 58–60, 253
- outcome switching, 110–12
- preprints, 221
- publication bias, 94
- replication crisis, 33, 36–41, 42
- Reuben affair (2009), 75
- spin and, 159–60
- statistical power and, 137
- Summerlin affair (1974), 54
- Wakefield affair (1998–2010), 76–9
- biomedical papers, 38, 70, 104, 137, 177
- bird flu, 217
- bispectral index monitor, 39, 41
- black holes, 239
- Black Lives Matter, 223
- blinding, 110, 134, 143, 178, 249
- blotting, 57–8, 60, 61, 69
- BMJ, The*, 77, 155
- Boldt, Joachim, 76
- books, 152–7
- Borges, Jorge Luis, 106
- Boulez, Pierre, 234, 235
- Boyle, Robert, 23, 24, 38
- brain imaging, 33, 137, 160
- Brass Eye*, vii
- British Medical Journal*, 77
- Brock, Jon, 163
- bronchoscopy, 52
- Broockman, David, 64–5
- Brown, Nick, 127
- Bush, George Walker, 98
- business studies, 109
- BuzzFeed News*, 102

- California Walnut Commission, 167
- California wildfires (2017), 175
- Canada, 124
- cancer
 - cell lines, 131–2, 143
 - collaborative projects, 218
 - faecal transplants, 161
 - food and, 164, 166
 - publication bias and, 94, 95
 - replication crisis and, 37, 38, 42
 - sleep and, 155–6
 - spin and, 159
- candidate genes, 139–42, 218
- carbon-based transistors, 71

Cardiff University, 149–51
cardiovascular disease, 94, 114, 167
Carlisle, John, 129–30, 168
Carlsmith, James, 128–9
Carney, Dana, 103–4
cash-for-publication schemes, 177–8, 182
cataracts, 47
Cell, 158, 162
cell lines, 131–3, 143
Cell Transplantation, 192
Center for Open Science, 233
CERN (*Conseil Européen pour la Recherche Nucléaire*), 89, 170
chi-squared tests, 87
childbirth, 39
China
 cash-for-publication schemes, 177–8, 182
 cell line mix-ups in, 131–2
 Great Famine (1959–1961), 244
 misconduct cases in, 69, 70
 randomisation fraud in, 159
chrysalis effect, 110
Churchill, Winston, 96
churnalism, 151
Cifu, Adam, 39–40
citations, 75, 186–9, 191–2, 194, 240
clickbait, 176
climate change, 34, 91, 236, 243–4
cloning, 55–7
Clostridium difficile, 161
cochlear implants, 47
Cochrane Collaboration, 41
coercive citation, 191, 192, 240
coffee, 150
cognitive dissonance, 128–9
cognitive psychology, 181
cognitive tests, 139, 140, 141, 150
coin flipping, 100–101, 129
Colbert Report, The, 3
Cold War, 242
collaborative projects, 217–18, 227–8
colonic irrigation, 164
communality, 21, 130, 176, 216, 218, 235, 242
COMPare Trials, 111
COMT gene, 140
confidence interval, 206
conflict of interest, 113, 115–16, 251
Conservative Party, 29, 115, 119
conspicuous consumption, 94

Cooperative Campaign Analysis Project (CCAP), 64–5
'Coping with Chaos' (Stapel), 4
Cornell University, 1, 98, 102
coronavirus (COVID-19), 225
Corps of Engineers, 175
correlation versus causation, 150, 156, 213, 251
corticosteroids, 139
Cotton, Charles Caleb, 123
creationism, 244
Crowe, Russell, 51
Csiszar, Alex, 243
Cuddy, Amy, 28–9, 32, 103, 104
CV (curriculum vitae), 180, 183, 194, 226, 229
cyber-bullying, 210
cystic fibrosis, 239

Daily Mail, 78
Daily Telegraph, 170
'Darwin Memorial, The' (Huxley), 81
Darwin, Charles, 21, 177, 182
Das, Dipak, 165
datasets, 35, 210
 fraudulent, 62–5, 86
 observational, 150
 publication bias, 86–112
Davies, Phil, 191–2
Dawkins, Richard, 22, 23, 90
De Niro, Robert, 78
De Vries, Ymkje Anna, 199–201
debt-to-GDP ratio, 124, 125
Deer, Brian, 77–8
democratic peace theory, 123
Denmark, 124
Department of Agriculture, US, 243
depression, 117, 140, 141
desk rejections, 18
Deutsche Bank, 221
disabilities, 47
discontinuous mind, 90
disinterestedness, 21, 85, 119, 176, 178, 227, 235
DNA (deoxyribonucleic acid), 58, 140, 141, 145–6, 217
domestication syndrome, 33
doveryai, no proveryai, 242
Duarte, José, 116, 118
Duke University, 240
duloxetine, 184
Dutch Golden Age, 196
Dutch Organisation for Scientific Research, 205

grapes, 165
 Great Recession (2007–9), 123
 Great Red Spot of Jupiter, 16
 Green, Donald, 65
 Gross Domestic Product (GDP), 124
 Gross, Charles, 69
 ground-breaking results, 57, 149, 157, 188, 204, 226
 groupthink, 114–15
 ‘Growth in a Time of Debt’ (Reinhart and Rogoff), 123, 233
 growth mindset, 152–4
 Guzey, Alexey, 155–6
 gynaecology, 159

h-index, 187–9, 192, 227, 228
 H5N1, 217
 Haldane, John Burdon Sanderson, 63
 Hankins, Matthew, 159
 HARKing, 99
 Harris, Sidney, 36
 Harvard University, 28, 223
 headache pills, 135, 138
 heart attacks, 40, 94, 167
 heart disease, 161, 213
 Heathers, James, 127, 150
 height, 62, 86–7, 88, 91, 135, 136, 138
 Heilongjiang University, 182
 Heino, Matti, 128
Henry IV (Shakespeare), 239
 Higgs Boson, 89
 Hirsch, Jorge, 187
 HIV (human immunodeficiency viruses), 239, 244
 homosexuality, 64, 75, 239
 Hong Kong, 53
 Hooke, Robert, 16
 Hossenfelder, Sabine, 110
 Houston, Texas, 223
 Hume, David, 13
 Huxley, Thomas Henry, 81
 Hwang, Woo-Suk, 55–7, 60, 74, 202
 hydroxyethyl starch, 76
 hype, 145–72, 196, 200, 219, 226

- arsenic life affair (2010), 145–8
- books, 152–7
- correlation versus causation, 150, 156, 213, 251
- cross-species leap, 149
- language and, 157–8
- microbiome studies, 161–4
- news stories, 150, 151

- nutrition and, 164–9
- press releases, 146–51
- spin, 158–60, 200
- unwarranted advice, 149

hypotheses, 16

Ig Nobel Prize, 98

images, fraudulent, 54–6, 60–61, 69–70, 202

impact factor, 190–92, 231

India, 55, 69, 70, 244–5

insomnia, 139

International Journal of Advanced Computer Technology, 185

Ioannidis, John, 42–3, 166, 207

IQ tests, 139, 140, 141, 150

Iraq War (2003–11), 30

Italy, 178, 244

Japan, 58–60, 70, 75, 253

John, Elton, 51

Journal of Cognitive Education and Psychology, 190

Journal of Environmental Quality, 243

Journal of Negative Results in Biomedicine, 204

Journal of Personality and Social Psychology, 3, 205

journals, 1, 15–21, 72

- conflict of interest disclosure, 113, 115–16
- fraud and, 60–61, 79
- hype and, 146–7
- impact factor, 190–92, 231
- language in, 157–8

journals (*Continued*)

- mega-journals, 205
- negligence and, 126, 129, 132, 138
- Open Science and, 218–19, 230
- peer review, *see* peer review
- predatory journals, 184–5, 194
- preprints, 221–5
- publication bias, 86–112, 116, 165, 196, 199–200, 204
- rent-seeking, 221
- replication studies, 205–6
- retraction, 61, 63, 66–9, 74, 75–6
- salami slicing, 182–4, 229, 230
- subscription fees, 218, 220–21

Jupiter, 16

Kahneman, Daniel, 25–8, 32, 247

Kalla, Joshua, 64–5

Karolinska Institute, 49, 50, 51, 52, 53, 202

Krasnodar, Russia, 49–50

This ebook is copyright material and must not be copied, reproduced, transferred, distributed, leased, licensed or publicly performed or used in any way except as specifically permitted in writing by the publishers, as allowed under the terms and conditions under which it was purchased or as strictly permitted by applicable copyright law. Any unauthorized distribution or use of this text may be a direct infringement of the author's and publisher's rights and those responsible may be liable in law accordingly.

VINTAGE
20 Vauxhall Bridge Road,
London SW1V 2SA

The Bodley Head is part of the Penguin Random House group of companies whose addresses can be found at global.penguinrandomhouse.com.



Penguin
Random House
UK

Copyright © Stuart Ritchie 2020
Cover Image © Alamy

Stuart Ritchie has asserted his right to be identified as the author of this Work in accordance with the Copyright, Designs and Patents Act 1988

First published in the UK by The Bodley Head in 2020

penguin.co.uk/vintage

A CIP catalogue record for this book is available from the British Library

ISBN 9781473564251