

PREFACE BY PETER WOOD

The study you have before you is an examination of the use and abuse of statistics in the sciences. Its natural audience is members of the scientific community who use statistics in their professional research. We hope, however, to reach a broader audience of intelligent readers who recognize the importance to our society of maintaining integrity in the sciences.

Statistics, of course, is not an inviting topic for most people. If we had set out with the purpose of finding a topic less likely to attract broad public attention, a study of statistical methods might well have been the first choice. It would have come in ahead of a treatise on trilobites or a rumination on rust. I know that because I have before me popular books on trilobites and rust: copies of Riccardo Levi-Setti's *Trilobites* and Jonathan Waldman's *Rust: The Longest War* on my bookshelf. Both books are, in fact, fascinating for the non-specialist reader.

Efforts to interest general readers in statistics are not rare, though it is hard to think of many successful examples. Perhaps the most successful was Darrell Huff's 1954 semi-classic, *How to Lie with Statistics*, which is still in print and has sold more than 1.5 million copies in English. That success was not entirely due to a desire on the part of readers to sharpen their mendacity. Huff's short introduction to common statistical errors became a widely assigned textbook in introductory statistics courses.

The challenge for the National Association of Scholars in putting together this report was to address in a serious way the audience of statistically literate scientists while also reaching out to readers who might quail at the mention of p-values and the appearance of sentences which include symbolic statements such as defining "statistical significance as $p < .01$ rather than as $p < .05$."

This preface is intended mainly for those general readers. It explains why the topic is important and it includes no further mention of p-values.

Disinterested Inquiry and Its Opponents

The National Association of Scholars (NAS) has long been interested in the politicization of science. We have also long been interested in the search for truth—but mainly as it pertains to the humanities and social sciences. The irreproducibility crisis brings together our two long-time interests, because the inability of science to discern truth properly and its politicization go hand in hand.

The NAS was founded in 1987 to defend the vigorous liberal arts tradition of disciplined intellectual inquiry. The need for such a defense had become increasingly apparent in the previous decade and is benchmarked by the publication of Allan Bloom's *The Closing of the American Mind* in January

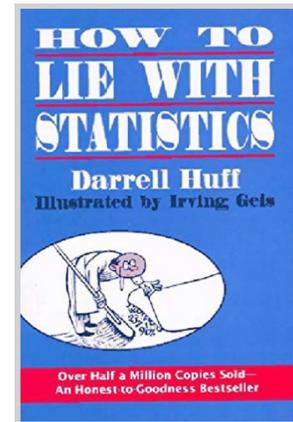


Figure 1: *How To Lie With Statistics* by Darrell Huff

1987. The founding of the NAS and the publication of Bloom’s book were coincident but unrelated except that both were responses to a deep shift in the temperament of American higher education. An older ideal of disinterested pursuit of truth was giving way to views that there was no such thing. *All* academic inquiry, according to this new view, served someone’s political interests, and “truth” itself had to be counted as a questionable concept.

The new, alternative view, was that college and universities should be places where fresh ideas untrammelled by hidden connections to the established structures of power in American society should have the chance to develop themselves. In practice this meant a hearty welcome to neo-Marxism, radical feminism, historicism, post-colonialism, deconstructionism, post-modernism, liberation theology, and a host of other ideologies. The common feature of these ideologies was their comprehensive hostility to the core traditions of the academy. Some of these doctrines have now faded from the scene, but the basic message—*out with disinterested inquiry, in with leftist political nostrums*—took hold and has become higher education’s new orthodoxy.

To some extent the natural sciences held themselves exempt from the epistemological and social revolution that was tearing the humanities (and the social sciences) apart. Most academic scientists believed that their disciplines were immune from the idea that facts are “socially constructed.” Physicists were disinclined to credit the claim that there could be a feminist, black, or gay physics. Astronomers were not enthusiastic about the concept that observation is inevitably a reflex of the power of the socially privileged.

The Pre-History of This Report

The report’s authors, David Randall and Christopher Welser, are gentle about the intertwining of the irreproducibility crisis, politicized groupthink among scientists, and advocacy-driven science. But the NAS wishes to emphasize how important the tie is between the purely scientific irreproducibility crisis and its political effects. Sloppy procedures don’t just allow for sloppy science. They allow, as opportunistic infections, politicized groupthink and advocacy-driven science. Above all, they allow for progressive skews and inhibitions on scientific research, especially in ideologically driven fields such as climate science, radiation biology, and social psychology (marriage law). Not all irreproducible research is progressive advocacy; not all progressive advocacy is irreproducible; but the intersection between the two is very large. The intersection between the two is a map of much that is wrong with modern science.

When the progressive left’s “long march through the university” began, the natural sciences believed they would be exempt, but the complacency of the scientific community was not total. Some scientists had already run into obstacles arising from the politicization of higher education. And soon after its founding, the NAS was drawn into this emerging debate. In the second issue of NAS’s journal, *Academic Questions*, published in Spring 1988, NAS ran two articles criticizing a report by the American Physical Society, that took strong exception to the quality of science in that report. One of the articles, written by Frederick Seitz, who was the former president of both the American Physical Society and the National Academy of Sciences, accused the Council of the



American Physical Society of issuing a statement based on the report that abandoned “all pretense to being based on scientific factors.” The report and the advocacy based on it (dealing with missile defense) were, in Seitz’s view, “political” in nature.

I cite this long-ago incident as part of the pedigree of this report, *The Irreproducibility Crisis*. In the years following the Seitz article, NAS took up a great variety of “academic questions.” The integrity of the sciences was seldom treated as among the most pressing matters, but it was regularly examined, and NAS’s apprehensions about misdirection in the sciences were growing. In 1992, Paul Gross contributed a keynote article, “On the *Gendering* of Science.” In 1993, Irving M. Klotz wrote on “‘Misconduct’ in Science,” taking issue with what he saw as an overly expansive definition of misconduct promoted by the National Academy of Sciences. Paul Gross and Norman Levitt presented a broader set of concerns in 1994, in “The Natural Sciences: Trouble Ahead? Yes.” Later that year, Albert S. Braverman and Brian Anziska wrote on “Challenges to Science and Authority in Contemporary Medical Education.” That same year NAS held a national conference on the state of the sciences. In 1995, NAS published a symposium based on the conference, “What Do the Natural Sciences Know and How Do They Know It?”



Figure 2: Frederick Seitz

For more than a decade NAS published a newsletter on the politicization of the sciences, and we have continued a small stream of articles on the topic, such as “Could Science Leave the University?” (2011) and “Short-Circuiting Peer-Review in Climate Science” (2014). When the American Association of University Professors published a brief report assailing the Trump administration as “anti-science,” (“National Security, the Assault on Science, and Academic Freedom,” December 2017), NAS responded with a three-part series, “Does Trump Threaten Science?” (To be clear, we are a non-partisan organization, interested in promoting open inquiry, not in advancing any political agenda.)

The Irreproducibility Crisis builds on this history of concern over the threats to scientific integrity, but it is also a departure. In this case, we are calling out a particular class of errors in contemporary science. Those errors are sometimes connected to the politicization of the sciences and scientific misconduct, but sometimes not. The reforms we call for would make for better science in the sense of limiting needless errors, but those reforms would also narrow the opportunities for sloppy political advocacy and damaging government edicts.

Threat Assessment

Over the thirty-one year span of NAS’s work, we have noted both the triumphs of contemporary science—and they are many—but also rising threats. Some of these threats are political or ideological. Some are, for lack of a better word, epistemic. The former include efforts to enforce an artificial “consensus” on various fields of inquiry, such as climate science. The ideological threats

also include the growing insistence that academic positions in the sciences be filled with candidates chosen partly on the basis of race and sex. These ideological impositions, however, are not the topic of *The Irreproducibility Crisis*.

This report deals with an epistemic problem, which is most visible in the large numbers of articles in reputable peer-reviewed journals in the sciences that have turned out to be invalid or highly questionable. Findings from experimental work or observational studies turn out, time and again, to be irreproducible. The high rates of irreproducibility are an ongoing scandal that rightly has upset a large portion of the scientific community. Estimates of what percentage of published articles present irreproducible results vary by discipline. Randall and Welser cite various studies, some of them truly alarming. A 2012 study, for example, aimed at reproducing the results of 53 landmark studies in hematology and oncology, but succeeded in replicating only six (11 percent) of those studies.

Irreproducibility can stem from several causes, chief among them fraud and incompetence. The two are not always easily distinguished, but *The Irreproducibility Crisis* deals mainly with the kinds of incompetence that mar the analysis of data and that lead to insupportable conclusions. Fraud, however, is also a factor to be weighed.

Outright Fraud

Actual fraud on the part of researchers appears to be a growing problem. Why do scientists take the risk of making things up when, over the long term, it is almost certain that the fraud will be detected? No doubt in some cases the researchers are engaged in wishful thinking. Even if their research does not support their hypothesis, they imagine the hypothesis will eventually be vindicated, and publishing a fictitious claim now will help sustain the research long enough to vindicate the original idea. Perhaps that is what happened in the recent notorious case of postdoc Oona Lönnstedt at Uppsala University. She and her supervisor, Peter Eklöv, published a paper in *Science* in June 2016, warning of the dangers of microplastic particles in the ocean. The microplastics, they reported, endangered fish. It turns out that Lönnstedt never performed the research that she and Eklöv reported.



Figure 3: Microplastics

The initial June 2016 article achieved worldwide attention and was heralded as the revelation of a previously unrecognized environmental catastrophe. When doubts about the research integrity began to emerge, Uppsala University investigated and found no evidence of misconduct. Critics kept pressing and the University responded with a second investigation that concluded in April 2017 and found both Lönnstedt and Eklöv guilty of misconduct. The university then appointed a

new Board for Investigation of Misconduct in Research. In December 2017 the Board announced its findings: Lönnstedt had intentionally fabricated her data and Eklöv had failed to check that she had actually carried out her research as described.

The microplastics case illustrates intentional scientific fraud. Lönnstedt's motivations remain unknown, but the supposed findings reported in the *Science* article plainly turned her into an environmentalist celebrity. Fear of the supposedly dire consequences of microplastic pollution had already led to the U.S. banning plastic microbeads in personal care products. The UK was holding a parliamentary hearing on the same topic when the *Science* article appeared. Microplastic pollution was becoming a popular cause despite thin evidence that the particles were dangerous. Lönnstedt's contribution was to supply the evidence.

In this case, the fraud was suspected early on and the whistleblowers stuck with their accusations long enough to get past the early dismissals of their concerns. That kind of self-correction in the sciences is highly welcome but hardly reliable. Sometimes highly questionable declarations made in the name of science remain un-retracted and ostensibly unrefuted despite strong evidence against them. For example, Edward Calabrese in the Winter 2017 issue of *Academic Questions* recounts the knowing deception by Nobel physicist Hermann J. Muller, who promoted what is called the “linear no-threshold” (LNT) dose response model for radiation's harmful effects. That meant, in layman's terms, that radiation at any level is dangerous. Muller had seen convincing evidence that the LNT model was false—that there are indeed thresholds below which radiation is not dangerous—but he used his 1946 Nobel Prize Lecture to insist that the LNT model be adopted. Calabrese writes that Muller was “deliberately deceptive.”

It was a consequential deception. In 1956 the National Academy of Sciences Committees on Biological Effects of Atomic Radiation (BEAR) recommended that the U.S. adopt the LNT standard. BEAR, like Muller, misrepresented the research record, apparently on the grounds that the public needed a simple heuristic and the actual, more complicated reality would only confuse people. The U.S. Government adopted the LNT standard in evaluating risks from radiation and other hazards. Calabrese and others who have pointed out the scientific fraud on which this regulatory apparatus rests have been brushed aside and the journal *Science*, which published the BEAR report, has declined to review that decision.

Which is to say that if a deception goes deep enough or lasts long enough, the scientific establishment may simply let it lie. The more this happens, presumably the more it emboldens other researchers to gamble that they may also get away with making up data or ignoring contradictory evidence.

Renovation

Incompetence and fraud together create a borderland of confusion in the sciences. Articles in prestigious journals appear to speak with authority on matters that only a small number of readers can assess critically. Non-specialists generally are left to trust that what purports to be a contribution to human knowledge has been scrutinized by capable people and found trustworthy. Only we

now know that a very significant percentage of such reports are not to be trusted. What passes as “knowledge” is in fact fiction. And the existence of so many fictions in the guise of science gives further fuel to those who seek to politicize the sciences. The Lönnstedt and Muller cases exemplify not just scientific fraud, but also efforts to advance political agendas. All of the forms of intellectual decline in the sciences thus tend to converge. The politicization of science lowers standards, and lower standards invite further politicization.

The NAS wants to foster among scientists the old ethic of seeking out truth by sticking with procedures that rigorously sift and winnow what scientific experiment can say confidently from what it cannot. We want science to seek out truth rather than to engage in politicized advocacy. We want science to do this as the rule and not as the exception. This is why we call for these systemic reforms.

The NAS also wants to banish the calumny of progressive advocates, that anyone who criticizes their political agenda is ‘anti-science.’ This was always hollow rhetoric, but the irreproducibility crisis reveals that it is precisely the reverse of the situation. The progressive advocates, deeply invested in the sloppy procedures, the politicized groupthink, and the too-frequent outright fraud, are the anti-science party. The banner of good science—disinterested, seeking the truth, reproducible—is ours, not theirs.

We are willing to put this contention to the experiment. We call for all scientists to submit their science to the new standards of reproducibility—and we will gladly see what truths we learn and what falsehoods we will unlearn.

For all that, *The Irreproducibility Crisis* deals with only part of a larger problem. Scientists are only human and are prey to the same temptations as anyone else. To the extent that American higher education has become dominated by ideologies that scoff at traditional ethical boundaries and promote an aggressive win-at-all-costs mentality, reforming the technical and analytic side of science will go only so far towards restoring the integrity of scientific inquiry. We need a more comprehensive reform of the university that will instill in students a lifelong fidelity to the truth. This report, therefore, is just one step towards the necessary renovation of American higher education. The credibility of the natural sciences is eroding. Let’s stop that erosion and then see whether the sciences can, in turn, teach the rest of the university how to extract itself from the quicksand of political advocacy.