“Should We Do Less Science?” is not a question that one hears being asked very often. One almost wonders if it is somehow socially or intellectually embarrassing to raise it. Partly, I suppose, that’s because the correct answer—“no,” of course—is so obvious that asking the question can only be viewed as some sort of mischievous provocation.

That the answer is so obvious draws mostly off of two different sources of reasoning, one philosophical and the other instrumental. The philosophical argument against doing less research reflects a commitment to the benefits for democratic societies of free and open inquiry, and the complementary idea that any attempt to limit the horizons of inquiry is inevitably subject to political abuse. As Roger Shattuck observes in his 1996 book *Forbidden Knowledge*, “Today, the principle of open knowledge and the free circulation of all goods and ideas have established themselves so firmly in the West that any reservations on that score are usually seen as politically and intellectually reactionary.” [167]

The instrumental reason builds off the standard portrayal of scientific research as inherently unpredictable in its findings and applications (that’s why it’s research, after all). Therefore, the consequence of choosing to do less of it—either by restricting the total amount done, or restricting some particular line of inquiry—has to be counterproductive because it simply means that you’re restricting the range of knowledge creation that might someday contribute
to human benefit. The unpredictability claim is in turn bolstered by the oft-repeated assertion that science, perhaps uniquely among human activities, is self-correcting. Self-correction, through peer review, competition among ideas and scientists, methodological rigor, and so on, guarantees continual improvement in the reliability of our knowledge about the world--progress. Self-correction assures us that more science is inherently more empowering of our ability to know and act than less science.

The power of these arguments manifests quite tangibly. For the most part, both the recent and the long-term history of science is one of continual growth of the enterprise as a reflection of a strong social consensus that more can, on the whole, only and always be better than less. I think it also important to point out that, as with any other societal activity that depends on state sponsorship, the growth in the number of those depending on that sponsorship in turn feeds a growing political interest in still more growth. So a nice complementarity exists between the elegance of the philosophical and instrumental arguments, and the political economy of the science enterprise itself. In the U.S., this confluence of rationale and politics remains powerful: When President Trump tried to significantly cut government funding for science as part of his overall assault on government programs of all types, our increasingly conservative, generally Trump-supporting Republican Congress refused to go along, and they in fact raised science funding to its highest levels ever.

But in a world where science is not just an expression of free inquiry, but also an expensive activity that is largely made possible by money provided by governments and corporations, we are constantly making choices about what science to do more of, and what to do (relatively or absolutely) less of. We do a lot more molecular genetics than ecosystems science or insect taxonomy; a lot more particle physics than seismology, or anthropology. From policy makers deciding on R&D funding allocations at the national level, to scientists participating in peer review of grant proposals, science policy processes are continually deciding that it would be better to do less of one kind of science and more of another.
And of course there are a few areas of science or avenues of scientific practice where explicit choices are made to actually forbid the pursuit of knowledge. Today, governments don’t support or permit research that would involve nonconsenting human subjects, and they often restrict research that causes undue suffering to certain species of animals, or that would violate laws about environmental protection or national security. These are rules about doing less science that pretty much everyone agrees with, and they tell us that there actually are moral and legal grounds for doing less science that democracies find unproblematic. Occasionally such decisions are controversial. For example, in 1993, when the U.S. Congress killed off the planned Superconducting Supercollider, and in so doing chose to do a lot less high energy physics, the resulting howls of protest focused on the folly of limiting our exploration of the unknown. In 2001, when President George W. Bush placed significant restrictions on embryonic stem cell research—restrictions that already existed here in Germany of course—the howls of protest focused on how partisan politics was interfering with both the open pursuit of knowledge and the potential for science to do good.

Areas of science occasionally do come under attack on ethical grounds or due to concerns about undesirable social consequences. The philosopher Philip Kitcher, in his book *Science, Truth, and Democracy*, laid out an argument on behalf of restricting some lines of research that were likely to undermine the public good. Should we allow research on the genetic determinants of intelligence? Should we pursue advanced molecular genetic therapies for rare diseases of rich people when scourges of the poor remain uncured? Ought research on nanotechnology or synthetic biology be slowed down or stopped while we try to better understand the potential risks? In a spasm of collective guilt following the unleashing of atomic weapons on Japan at the end of World War II, some American physicists pushed for the US government to do less of the sorts of physics research that could lead to more weapons. But even Robert Oppenheimer, confessing in 1947 that “physicists have known sin,” [Shattuck, 176] also insisted: “If you are a scientist you believe that it is good to find out how the world works. that it is good to turn over to mankind at large the greatest possible power to control the world and to deal with it according to its lights and values.” [Shattuck, 184]
I’m not going to talk about these sorts of questions, which are largely addressed as matters of ethics and values. Instead, I want to focus on two aspects of today’s science enterprise that raise very practical questions about the consequences of doing more science rather than less. The first of these is a problem of quality, the second a problem of epistemology. It turns out that these problems are, perhaps not surprisingly, quite closely related, but for the moment I’ll keep them separate.

The problem of quality in science is now widely acknowledged, and is often termed a crisis in reproducibility or replicability—a framing which itself unwittingly reflects epistemological matters that I’ll get to later. That poor quality of published scientific results was a widespread problem became unavoidably public in the early years of this decade when scientists in the pharmaceutical industry found that they were only able to replicate in their laboratories a small percentage of academic research studies of the types they normally depend on to identify promising drugs. A broader investigation leading to a cover story in the *Economist* pointed to problems of reliability across a broad swath of scientific disciplines, and a recent book by the science journalist Richard Harris documented systemic problems of poor quality and lack of reliability across the biomedical sciences, and a book by the theoretical physicist Sabine Hossenfelder questioned the quality of much of the research done in her field. No summary of these problems is complete without mention of John Ioannidis, now at Stanford, who has done statistical analyses of fields ranging from genomics to nutrition to economics, and demonstrated systemic bias and lack of reliability across all of them. His now-famous 2005 article, “Why Most Published Research Findings are False,” provides a general explanation rooted in statistical theory for the rise of poor-quality science. The practical implications of these collective findings are profound and sweeping. As just one example, Ioannidis and colleagues note, in a 2017 paper, that “much published medical research is not reliable or is of

---

2 https://www.economist.com/leaders/2013/10/21/how-science-goes-wrong
3 https://www.basicbooks.com/titles/richard-harris/ rigor-mortis/9780465097913/
5 https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124
uncertain reliability, offers no benefit to patients, or is not useful to decision makers.” Think about that the next time you decide to go to the doctor.

The scientific community has allowed itself a certain degree of self-congratulation because of its willingness to acknowledge that these quality problems exist, and has framed this as “self-correction in science at work.” but the scale of efforts to address the problem seems utterly unequal to its magnitude and pervasiveness. Survey data show that many scientists these days agree that the incentive system in academic science, the hypercompetitive environment for funding and publication in high-prestige journals, and the bibliometrics tools used to quantify scientific performance, have created an environment that provides few systemic checks on quality, and which allows and even encourages poor quality science to quickly proliferate in the scientific literature. A factor that strongly contributes to this set of pathologies is the sheer size and diversity of the research enterprise, which makes quality control very tough to impose in any systematic way, and which provides a growing array of outlets—journals, websites, and so on—for moving data and information from the laboratory out into the world with little accountability for quality. As just one small example of this problem, a 2012 article in Science showing that an Alzheimer’s drug called bexarotene would reduce beta-amyloid plaque in mouse brains was reported in 2013 to be unreproducible, yet it has been cited more than 750 times, including 123 times since 2016. In this way, poor-quality research metastasizes through the published scientific literature, and distinguishing knowledge that is reliable from knowledge that is unreliable or false or simply meaningless becomes impossible. I should add that this general scenario—system dysfunction as a result of exponential system growth—was first anticipated more 60 years ago by the physicist and historian of science Derek de Solla Price. He called it “scientific doomsday.”

---

6 http://science.sciencemag.org/content/335/6075/1503
9 http://derekdesollaprice.org/little-science-big-science-full-text/
In the U.S., several efforts are being aimed at assessing and improving quality of science in certain fields, most notably the Center for Open Science, led by University of Virginia social psychologist Brian Nosek, and the Meta-Research Innovation Center at Stanford, led by Ioannidis and biomedical researcher Steve Goodman. Currently, the US National Academy of Sciences is conducting a study, at the request of the US Congress, on reproducibility and replicability in science. Such efforts view the quality problem as one of internal culture that will have to be solved through better training of scientists, through a realignment of professional incentives to reward quality not quantity, through more rigorous statistical practice in data analysis, better problem choice, journals that publish negative experimental results as readily as they publish positive ones, and so on.

These efforts are tiny compared to the potential scale and implications of the problem. One can reasonably say that significant portions of the multi-hundred-billion-dollar global public investment in scientific research are being compromised by the proliferation of poor-quality science. Unless one wants to end up arguing that, in effect, it’s okay to fund poor-quality science, then one ought to take seriously the need to figure out how to fund less of it. In the U.S., the current political economy and culture of academic science, at least, seem unable to address this need at the enterprise scale. Perhaps we need a temporary slow-down in research funding, matched by a significant national assessment and reform effort? Certainly less, but better, science would be preferable to more, but poor quality, science.

Yet viewing the quality problem as a problem to be solved through internal system reform seems to miss what may be the most vexing aspect of the quality problem. Let me pursue this with reference to the cold mouse problem.

Mice are commonly used to explore cancer behavior and to test new anti-cancer drugs. Since 2011, US National Research Council guidelines for standardized care and use of laboratory mice requires that mice be housed at temperatures of between 20°C and 26°C. Prior to 2011, the required temperature range was between 18°C and 24°C. In 2013, an article in PNAS
(Proceedings of the National Academy of Sciences)\textsuperscript{10} reported that housing laboratory mice at such temperatures significantly reduces tumor immunity, and encourages tumor formation, growth rate, and metastasis, in comparison to mice housed at warmer temperatures of 30°C-31°C.

While the article mentions that the effects of mouse housing temperature “has received little attention,” it does not explicitly make the point that this new variable should call into question the value of mouse model research on cancer that doesn’t consider temperature as a variable in both cancer behavior and cancer drug testing. Rather, the article treats the finding as science marching forward: “[A] better understanding of the physiological interactions between stress responses, thermoregulation, and immune regulation could reveal important new strategies for strengthening antitumor immunity.”

As a 2015 article in Nature Communications\textsuperscript{11} makes clear, however, this new understanding means that testing cancer drugs in cold mice “may not accurately predict which new therapies will be effective in the clinic.” In other words, up until just a few years ago, pre-clinical tests of new cancer therapies failed to account for what turns out to be a key variable in cancer biology in mice. Then, this past July, a study by data scientists at Elsevier\textsuperscript{12} showed that 94 percent of 133,000 papers reporting on tumor research in mice since 2000 did not report housing temperatures, and that neither this reporting frequency, nor the housing temperatures used for mice, changed in studies published after the 2013 and 2015 papers that first reported the problem.

Housing temperatures for mice? Who knew? Who knew this was a key variable in tumor formation and tumor immunity in mice? Who knew that housing temperatures should be reported as an experimental variable? In fact, wasn’t standardization of experimental temperatures an important way to ensure comparability across research projects?

\textsuperscript{10} http://www.pnas.org/content/110/50/20176
\textsuperscript{11} https://www.nature.com/articles/ncomms7426
\textsuperscript{12} https://ieeexplore.ieee.org/document/8258456
Here is where the problems of quality and epistemology intersect, or perhaps it’s better to say they collide. When mouse-tumor researchers weren’t aware of the importance of housing temperatures, they could hardly be held accountable for not reporting them. Their experiments may in fact have been reproducible if whoever was doing the reproducing happened to house their mice at the same temperatures as the original experiments. On the other hand, the fact of such experiments being reproducible may have said very little about whether results were actually useful as pre-clinical indicators of drug efficacy, because they did not incorporate the key variable of temperature. Given that about 90 percent or so of all cancer drug trials ultimately fail, would keeping mice at higher, less-cancer-friendly temperatures give rise to pre-clinical results that are more predictive of efficacy in humans? No one knows. And what is to be done with the 125,000 papers published since 2000 that use mice to study tumors without reporting on temperature?\textsuperscript{13} Sure, this may be called self-correction, but it is also the case that many of these papers now look like poor quality science, having ignored a key variable in the processes they are supposed to be shedding light on. And they will remain part of the knowledge base, to be cited, used to generate new hypotheses, identify new possible cancer drugs, and so on. Such difficulties call into question the very idea that scientific advance is cumulative.

There is always another cold-mouse problem at the intersection of quality and epistemology when systems are open, complex, and incompletely understood. But at least mouse-tumor research has two additional dimensions that make it potentially tractable. First, everyone would like to see better treatments, cures, and prevention for cancer. So the values that motivate science are broadly shared across society. Second, efficacy in treatment of cancer in humans therefore provides an end-point for assessing quality. This is a critically important point: ultimately, poor quality in cancer science will manifest as real-world failure in drug development. Remember, I mentioned earlier that it was actually the pharmaceutical industry, not academic researchers, that first blew the whistle on the quality problem in pre-clinical

\textsuperscript{13} Actually the number will be much greater, since the study only looked at papers in Elsevier-published journals.
biomedical science. That’s because drug company researchers, tired of high failure rates in clinical trials, decided that maybe the cause of the problem lay in the quality of the science that they were depending on to choose potential cancer drugs. When shared values converge around distinct endpoints for science, then at least you know what counts as success, so reasonably stable criteria of quality can be inferred.

Things really begin to get complicated when such convergence is not possible—when there is disagreement over values, and when there is no real-world end-point as a check on quality.

And that brings me to the tale of Chlorine 36.

The Chlorine 36 story comes out of the nearly 30 years of scientific effort to calculate the behavior of groundwater at the proposed Yucca Mountain nuclear waste repository site in Nevada. The flow of water through the site 300 meters underground can compromise site integrity both by contributing to corrosion of waste-containment vessels, and by transporting radionuclides from the site into the surrounding environment. The behavior of water at the repository site is thus a crucial element of the overall capacity to safely store waste over tens of thousands of years, and has been an object of intensive study.

A standard way of characterizing groundwater flow at Yucca Mountain has been in terms of “percolation flux,” or the amount of water moving through a unit area in a given time, usually reported as millimeters per year. Early estimates for Yucca Mountain repository rocks were based on the expert judgment of a couple of geologists who had spent many years doing field work in the region, and on this basis a 1983 US Geological Survey report estimated a percolation flux of 4.0 mm/year for the rocks at the repository site. A year later, another USGS report lowered that estimate to 1.0 mm/year. These were ballpark estimates provided by experts based on their assessment of geologic,
hydrologic, and climatic conditions in an area they knew well. Here the idea of quality would refer to craft-based scientific judgment. (Metlay, 2000).

Over the next decade, seven formal determinations of percolation flux were published as part of environmental and performance assessment activities for Yucca Mountain. These determinations were no longer the products of mere expert judgment. Rather, they were derived from laboratory tests on rock cores from the site, and from numerical models of water flow through rock. During this period of increasing scientific sophistication, the mean value for percolation flux estimates stabilized at between 0.02 and 1.0 mm/yr (Metlay, 2000).

Now this looks like the progress of science. A convergence of results carried out over multiple years, and reflecting multiple sophisticated and quantitative scientific methods, using both physical samples and mathematical models rather than mushy, irreproducible craft judgment. Not only that, but it turned out that these numbers were converging right where they needed to be—at or below 1.0 mm/yr, and thus consistent with performance assessment models that could meet the required safety standards for site behavior over the long-term (Metlay, 2000). Reproducibility in action.

In the early 1990s DOE began excavations to allow sampling from the rocks at the depth of the actual repository site. Groundwater along an exploratory tunnel about 300 meters under the surface was analyzed for trace chemicals at Los Alamos National Laboratory, and the results showed elevated levels of Chlorine 36. While this isotope occurs naturally in very small amounts in air, most Cl$^{36}$ was created by atmospheric nuclear weapons tests in the 1950s and 1960s. Yucca Mountain is very close to the Nevada weapons test site, and the discovery of elevated levels of bomb-pulse chlorine meant that water containing this radioisotope had managed to travel from the earth’s surface to a depth of 300 meters in a period of 40 years or so. This was much faster than the hydrological models of the repository area had predicted, and it challenged the well-replicated results of the previous decade. Yet there was no obvious way to integrate the isotopic findings into either lab tests or existing models. So DOE went back to expert judgment, and conducted an expert elicitation where seven experts derived subjective probabilities based on their assessment of available data and knowledge of the site—craft
skill again. These distributions were combined to yield a mean percolation flux of 10 mm/year, basically an order of magnitude higher than what the past decade of replicated experiments and models had determined. So the site was wetter than had previously been thought. This was a very important change, because it called into question the suitability of the site for long-term geological storage (Metlay, 2000).

Science marches on. Except when it doesn’t. Here’s what happened next. Efforts to replicate the Los Alamos isotopic analysis were conducted at Lawrence Livermore National Lab. This study determined that the elevated Cl$^{36}$ in the water samples analyzed by Los Alamos scientists was a result of faulty sample preparation and likely contamination. Lawrence Livermore scientists concluded there was no evidence for higher percolation flux. Then Los Alamos scientists repeated their analysis and replicated their earlier results showing that the original determination of elevated Cl$^{36}$ was a valid result after all. Then Lawrence Livermore replicated their own earlier results showing that it wasn’t.

Just to give you a sense of the analytical precision involved here, the bomb pulse Chlorine to atmospheric Chlorine ratios determined by Los Alamos scientists was on the order of $10^{-12}$. The Lawrence Livermore results were on the order of $10^{-13}$, more-or-less consistent with the ambient atmospheric ratio. This factor of one-in-a-billion seemed to translate into what could be the difference between a safe repository site and an unsafe one.

Enter a cold mouse. It turned out that some basic questions about sample preparation had not been recognized. For example, the Los Alamos scientists ground their rock samples more finely than the Livermore scientists, and the latter believed that this led to artificially elevated Cl$^{36}$ levels in the Los Alamos results. Who knew that the results of the radioisotope analysis would hinge on how finely you ground your sample? Well, it’s always something. And it looks like the methodological uncertainties in this work are greater than the precision necessary to distinguish the presence of Cl$^{36}$ in groundwater. The 2006 Geological Survey study\(^\text{15}\) summarizing these and other complications concluded that the discrepancies between the different studies “cannot be explained by presently available data.” [p. 59]

\(^{15}\) https://www.nrc.gov/docs/ML0915/ML091590561.pdf
A completely new, independent study\textsuperscript{16} by scientists at University of Nevada, Las Vegas was also unable to resolve the source of the discrepancies.

If only the Los Alamos scientists had ground their samples more coarsely! Then the Cl\textsuperscript{36} results might have replicated the early experimental and modeling results, and no one would have been the wiser!

Instead, we are left with uncertainties greater than they have ever been. The Cholrine discrepancy remains unresolved. Most recently, a 2013 hydrological modeling study\textsuperscript{17} concludes that percolation flux lies in a range between about 1.0 and 20.0 mm/yr. Thirteen billion dollars of research later, we are back to the same order-of-magnitude results as those that individual field scientists applying craft judgment came up with in the 1980s.

This thirteen billion dollars is a direct investment in the delusion that the complex scientific problems, and the deeply contentious politics, of Yucca Mountain, were separate domains that could be dealt with independently and sequentially. If the behavior of groundwater at the repository site were of merely academic interest, then the percolation flux value could have been settled at any point during the past three decades merely by not doing any more research, and no one would have cared one way or the other. It was only because of bitter political disagreements about the risks of Yucca Mountain that numbers like percolation flux were continually scrutinized by scientists and stakeholders. But percolation flux is an abstraction that cannot be directly measured, it must be inferred from experiments, models, and observations as part of the effort to provide predictive accuracy of site behavior for the next 10,000 years. What I am emphasizing here is one cannot understand what is going on scientifically without understanding what is going on politically. Standard notions of science—replicability, self-correction, independence, intersubjectivity, empiricism, convergence on reliability—do not tell the story.

\textsuperscript{16} https://digitalscholarship.unlv.edu/yucca_mtn_pubs/67/
\textsuperscript{17} https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013WR014122
Funtowicz and Ravetz’ concept of Post-Normal Science\textsuperscript{18} provides a conceptual framework for understanding the percolation flux story at Yucca Mountain. Post-Normal Science describes the metamorphosis that science undergoes when four conditions pertain: decision stakes are high; decisions urgently need to be made; facts are contested; and values are conflicting. I would add to this list: when there is no agreed upon end-point toward which scientific progress and quality can be assessed.

Post-normal science is born from the expectation that both uncertainty about complex system behavior, and disagreement about policy actions to be taken, can be reduced through more scientific research of the conventional type. But in the complex systems that are the object of investigation for post-normal science, this expectation is backwards. There are always multiple ways to characterize and study a problem, there are always facts on one side of a question and on another, there is always another cold mouse problem or a Cl\textsuperscript{36} problem waiting to be unearthed. Perhaps most importantly, there are always competing sets of actors who draw from the science what they need to reinforce pre-existing self-consistent views of the world as they understand it and as they want it to become.

The cultural theory of Mary Douglas\textsuperscript{19} and her disciples provides a useful adjunct to the post-normal science framework, a different dimension, the dimension of plural rationalities. For Yucca Mountain, we can understand the government agencies that were responsible for certifying the safety of the site as embodying a hierarchical, rules-based rationality that depends on normal science to justify its actions. We can understand the role of the nuclear power operators in terms of individualistic, market-based rationality that is basically allied with the government because they want to pass off the responsibility for the waste to the government. For these two ways of thinking about the world, 30 years of research have shown that the uncertainties have been reduced as much as possible, and that action should be taken. As one government scientist said: “The information we have is that there’s basically nothing

\textsuperscript{18} For an early presentation of the concept, see: https://www.uu.nl/wetfilos/wetfil10/sprekers/Funtowicz_Ravetz_Futures_1993.pdf
\textsuperscript{19} For example, see: https://www.amazon.com/Divided-We-Stand-Re-Defining-Technology/dp/081221319X
wrong with that site, and you’re never going to find a better site.” Yet we can understand continued fierce opposition to the site by Nevada residents and environmental groups as expressing an egalitarian and communitarian rationality that understands ongoing uncertainties as the strongest reason not to take action. “Everything that the Department of Energy has done since 1987 is to put political science ahead of earth science,” says one Nevada state official.

The combination of post-normal science and plural rationalities characterizes many problems against which science has been hurled, often for decades or more: the efficacy of mammograms and PSA tests, the effects of various nutrients, foods and diets on health and life expectancy, the carcinogenicity of a wide range of chemicals, the best ways to reduce obesity, the sustainability of many natural resources, the importance of genetically modified foods for global food security, the role of globalized markets in causing job loss, the risks of nuclear power, the rate of animal extinction, the effectiveness of alternative ways to teach reading and math . . . . and on and on.

For almost all of these problems, and many others like them, scientific information on all sides of the matter is available. The internet can instantly dish up peer-reviewed validation of pretty much any position one would like to take on such matters. In a critique of nutrition science published just a few weeks ago, John Ioannidis observes: “Almost all nutritional variables are correlated with one another; thus, if one variable is causally related to health outcomes, many other variables will also yield significant associations in large enough data sets. With more research involving big data, almost all nutritional variables will be associated with almost all outcomes.”

This, truly, is science for the people. The persistent claim that we are in a post-truth world seems utterly to miss what’s going on underneath all the politics: truth has been knocked off its pedestal and is now available to everyone to support mutually contradictory beliefs across an extraordinarily broad range of social and political concerns. We’re awash in truth, promiscuously provided by what everyone persists in calling science. To put it in obnoxiously American terms, science has badly cheapened the value of its brand name through relentless
marketing and hype; like hamburgers and knock-off Rolexes, science is everywhere, available to everyone. Science is bling.

Now: We all know, more-or-less, what we mean to convey when we invoke science. We are conjuring up the power of modernity and Enlightenment, we are invoking Galileo against the Church, rationality against superstition, controlled experiment against folk wisdom, climate models against Big Oil. We are saying that science alone provides reliable knowledge, that it will allow us to distinguish between fact and belief, that we will achieve greater clarity about the matters under consideration, and that findings, however imperfect or incomplete, must still be privileged over assertions about the world that lack science’s imprimatur and authority. We are saying that scientific findings must take precedence over non-scientific assertions in guiding our actions. Above all, we mean to distance assertions backed by science from our own, or anyone else’s, subjective take on the world. Even if we don’t always believe all that, it’s what we mean when we say “the science shows this” and “the research suggests that.”

Post-normal science is a different thing from what we mean when we talk about science. Post-normal science is not science that’s very difficult, or science that’s uncertain, or science that needs more funding. Post-normal science is the activity that occurs when the tools, practices, and expectations of normal science are applied in the dual contexts of system complexity and political disagreement. In post-normal science, problems of quality blend into problems of politics, because decisions that scientists always must make, about experimental design, model assumptions and parameters, data choice, statistical significance, and so on, cannot be kept separate from assumptions about how the world works, or should work. In normal science, the Chlorine 36 problem is a confounding factor and a further source of uncertainty, an opportunity for more research. In post-normal science, the Chlorine 36 problem is another opportunity for competing interests to advance their positions.

Should we do less science? The answer is: We already are. But we remain in denial about this.

The benefits of our denial are clear. For decision makers, it’s always easier to call for more science to clarify options and to support pre-existing commitments than to make choices that
are politically risky. Advocates on various sides of these debates need science so they can re-
assert their own commitment to rationality and evidence, and portray their opponents as ir-
rational. And scientists, of course, get to keep doing what they like to do. The competing and con-
flicting values that underlie disagreement are hidden behind arguments about science that are endlessly fed with more cold mice and chlorine isotopes.

The ideals of truth and science are meanwhile cheapened, because everyone knows that the science is on their side, and the other side is just throwing junk science around to advance their own interests. Both skepticism and credulity are in this way encouraged. I don’t think it’s at all unreasonable to suggest that these conditions have helped create opportunities for demagogues like Donald Trump to gain credibility and power.

Yet I do think there’s actually something useful and rather simple to be done here. People who care about the quality and legitimacy of science could start insisting at every chance that science conducted and invoked in the post-normal context is not science. Post-normal science is easy to spot. When experts continue to disagree; when advocates continue to use science to advance value-based agendas and to accuse those they disagree with of misusing science; when decision makers don’t take action on urgent issues but call for more research; when action means that there will be winners and losers; when the quality of the science cannot be measured against any agreed-upon end-point—then, no matter how sophisticated the math or complex the scientific instruments, no matter how pure of motive and careful of method the scientists, it’s NOT science, and we should all say so.

Is there an clear and unambiguous demarcation between normal and post-normal science? Of course not, but that goes for almost any categorical boundary between related but different things. We can argue about research activities that fall into the gray area, but at least such arguments would contribute to the larger idea that on either side of the gray area are two fundamentally different sorts of inquiry.

Post-normal science is not science, and we should say so—but this does not mean that inquiry in the post-normal context is not worthwhile, is not a potentially valuable adjunct to
democratic problem-solving. The insistence that science is a precondition for rational action in the face of political disagreement ends up stifling the democratic imagination, because decision makers can just wait around—forever—for the science relating to current options to get settled, which it never does. Or they can impose a solution and avoid accountability by claiming that the science made them do it. In the post-normal context, the opposite stipulation pertains. Given conditions of political disagreement, urgency, system openness, and so on, problems and solution paths have to be discussed and agreed upon as a part of the process of knowledge creation. Inquiry in the post-normal context is thus an invitation to democratic creativity. At the end of a long talk, this is a point to be fleshed out elsewhere, but I will simply suggest that there are many encouraging instances, typically at the margins of the mainstream research system, that are taking what amounts to a post-normal approach to science. Examples include the Toxics Use Reduction Act in Massachusetts, the Regional Integrated Science Assessment program at the National Oceanographic and Atmospheric Administration, and Sweden’s approach to nuclear waste disposal.

But what, then, of quality? If quality in normal science is supposedly assessed through familiar, although sometimes difficult-to-apply, standards of methodological rigor, cumulative achievement of intersubjective knowledge, and progress toward objective, agreed-upon endpoints, quality in post-normal science is perhaps best seen in terms of fitness-for-purpose. In the post-normal context, quality is as much or even more a matter of appropriate institutional arrangements and expectations than of precision, accuracy, reproducibility, and so on.

The links between the problem of quality in science and the question of post-normal science thus need to be made much more explicit, because I doubt we can make much progress solving either problem in isolation from the other. John Ioannidis and others who are concerned about problems of scientific quality are focusing on reform of normal science within the science system itself. An unavoidable implication of their work is that a lot of published science is not just worthless but wrong. Most of this wrong science is affirmatively wrong—that is, the things it says are true are in fact either false or indeterminate. They are false positives. This is unsurprising. Scientists want to discover things, of course, so they search for, and get rewarded
for publishing, positive findings, not negative ones. But because quality problems manifest as false positives, if we could somehow magically turn poor quality published studies into high quality studies, the result would not be new high quality knowledge, but a mountain of mostly negative findings. It would be as if we had done less science to begin with. So in this way the quality problem yields the same answer to my question as the epistemological problem: we already are doing less science.

Indeed, in the recently published paper about nutrition science that I mentioned earlier, Ioannidis suggests that “Resources for some of these studies could have been better spent on unambiguous, directly manageable threats to health such as smoking, lack of exercise [or] air pollution.” This begins to sound a lot like what I’ve been suggesting for the post-normal context: recognizing the crisis in scientific quality not only as a call for reform, but as a reason to creatively search for alternatives, for better, perhaps more direct ways to define and address problems that have refused to yield to science-led solutions.

Science has been an immeasurably valuable gift from humans to themselves. But in a world increasingly challenged to manage the complex conditions of globalization, development, sustainability, and continuous technological change, science is not the solution to everything, and sometimes is even part of the problem. Looking this problem squarely on means also being willing to question some of our most cherished beliefs about rationality, action, and human choice. But being willing to question cherished beliefs is perhaps above all the essence of the scientific attitude. Today, science is faced with major threats to its legitimacy and social value. A scientific attitude about science itself may help. But I think some common sense can help, too. We all know that for many important things in life, excess can lead to complacency, corruption, and even ruin. We are already doing a lot less science than we believe. If we can open ourselves up to that idea, then we might find that what remains is all the more precious.