

# Criminalizing Science Fraud

And other ideas for promoting scientific integrity.

By Glenn Harlan Reynolds

Should we criminalize fraud in science? Some people are saying so.

As Chris Said [writes](#):

*In 2006, Sylvain Lesné published an influential Nature paper showing how amyloid oligomers could cause Alzheimer's disease. With over 2,300 citations, the study was the 4th most cited paper in Alzheimer's basic research since 2006, helping spur up to \$287 million of research into the oligomer hypothesis, according to the NIH.*

*Sixteen years later, Science reported that key images of the paper were faked, almost certainly by Lesné himself, and all co-authors except him have agreed to retract the paper. The oligomer hypothesis has failed every clinical trial.*

*Lesné's alleged misconduct misled a field for over a decade. We don't know how much it has delayed an eventual treatment for Alzheimer's, and it was not the only paper supporting the oligomer hypothesis. But if it delayed a successful treatment by just 1 year, I estimate that it would have caused the loss of 36 million QALYs (Quality Adjusted Life Years), which is more than the QALYs lost by Americans in World War II. (See my notebook for an explanation.)*

*Lesné is not alone. This year we learned of rampant image manipulation at Harvard's Dana-Farber Cancer Institute, including in multiple papers published by the institute's CEO and COO. So far 6 papers have been retracted and 31 corrected. The 6 retracted papers alone have 1,400 citations and have surely polluted the field and slowed down progress. If they*

delayed a successful cancer drug by just 1 year, I estimate they would have caused the loss of 15 million QALYs, or twice the number of QALYs lost by Americans in World War I.

A QALY is a Quality Adjusted Life Year, a measure of health / longevity benefits.

The point is that scientific fraud can have serious consequences. (Especially when, as is often the case, the fraudsters attain a sufficiently elevated position to starve competing ideas of grant money.)

Said notes that these aren't the only examples:

Alleged research misconduct	Citations	Hypothetical delay in an effective treatment caused by misconduct	Potential QALYs lost	Outcome
<a href="#">Sylvain Lesné (Alzheimer's)</a>	2,300 <sup>1</sup>	1 year	36 Million	Lesné still at UMN
<a href="#">Berislav Zlokovic (Alzheimer's)</a>	11,500 <sup>2</sup>	0.5 years	18 Million	Ongoing investigation
<a href="#">Harvard's Dana-Farber Cancer Institute (Cancer)</a>	1,400 <sup>3</sup>	1 year	15 Million	Ongoing investigation
<a href="#">Carlo Croce (Cancer)</a>	3,374 <sup>4</sup>	0.5 years	8 Million	Croce still at OSU

**Table 1.** Examples of alleged research misconduct and the hypothetical delay they may have caused to the development of effective treatments. Citation counts are outlined in [this spreadsheet](#) and QALY calculations based on [this notebook](#).

Disclaimer: Views expressed are my own.

<sup>1</sup>Citations of the [Nature paper](#), not including the [19 other papers](#) suspected to contain doctored images.

<sup>2</sup>Citations of [35 papers](#) where whistleblowers have raised concerns.

<sup>3</sup>Citations of the 4 papers publicly listed as retracted so far.

<sup>4</sup>Citations of [15 retracted papers](#).

You can dispute these death numbers, which are admittedly hypothetical, but you can cut them by a factor of ten, or even a hundred, and they're still big numbers of deaths caused by fraud. (Remember, we're not talking about mistakes or errors here, but fraud.)

We certainly prosecute people for much less dramatic frauds all the time, frauds that have much less impact.

So it makes sense to punish science fraud, which has the potential to do tremendous harm. But.

There are a few problems. First, previous experience with trying to police science fraud suggests that our authorities aren't very good at it. As I [wrote](#) quite a while ago – originally in a chapter on the ethics book I wrote with Peter Morgan, *The Appearance of Impropriety* – efforts at policing science fraud in the 1980s and 1990s were a disaster.

As I wrote, in response to hearings on science fraud chaired by then-Senator Albert Gore, Jr., the National Institutes of Health created an office of Scientific Integrity (later renamed the Office of Research Integrity), which expanded its mission from punishing scientific fraud to a broader, less-focused mission of policing “scientific misconduct.”

The results were awful. Malcolm Gladwell summarized its work in an article for the *Washington Post* entitled “The Fraud Fraud,” concluding that up close, the sins of scientists identified by ORI looked minor.

Rameshwar Sharma, a researcher at the Cleveland Clinic, made a typo in a grant application, invoking in one place a different protein than the one he was actually researching. The application was denied, and nobody was fooled: It was quite clear from context which protein he was actually researching, since that was the one all the data concerned.

An anonymous accuser pointed him out to the ORI, which chose to interpret his typo as deliberate fraud aimed at securing a federal grant. This produced a three-and-a-half-year nightmare for Sharma, one boosted by Rep. John Dingell (D-MI) who, along with his staff, kept the pressure on the case up. In the course of sparring with Sharma's boss, Bernadine Healy, Dingell claimed that Sharma was “ultimately found to have falsified his grant application.”

This was false; the matter was under appeal and there had been

no ultimate finding. Sharma's career was devastated. His lab was shut down, he would up having to take his kids out of college, and for a while he was working in an unsalaried position at an optometry college while living in a dormitory. Dingell, in fact, was guilty of what he (falsely) accused Sharma of – deliberately overstating his case in order to draw attention. (Ironically, Dingell made this accusation in an article in the *New England Journal of Medicine*, so he was even guilty of making a false statement in a scientific journal).

An appeals panel at the NIH ultimately vindicated Sharma – that is, he was “ultimately found” *not* to have done anything wrong. The panel said a careless error, not misconduct, was involved, and that ORI's findings were not supported and the proposed punishment of Sharma was not justified. In spite of this contrary finding, Dingell's office, rather than retracting its statements, doubled down.

And all the leading cases brought by ORI had similar problems. Margit Hamosh, a researcher at Georgetown, was charged with “anticipatory research” based on a single ambiguous statement in a 20,000 word grant application that ORI chose to interpret, somewhat illogically, as a deliberate misrepresentation.

Similar behavior accompanied the investigations of researchers Mika Popvic and Robert Gallo, and Thereza Imanishi-Kari and David Baltimore. (At one point, referring to Gallo and Baltimore, a Dingell staffer [boasted](#) “we have taken out the two biggest names in science.”) All the investigations collapsed.

So the track record here is poor, to say the least, and heavily marked by politics. Indeed, ironically enough, the lawyers – usually cordially detested by scientists – were the only group to come out ahead in this whole business. Bernadine Healy said that the ORI process had become a “star chamber” in which scientists behind closed doors used its

investigations to take down their rivals. Robert Gallo said “if it weren’t for the lawyers we’d all be dead.”

Would we do better today? I doubt it.

We’ve seen the way that the scientific establishment went after purveyors of entirely scientific doubts about Covid policy, cancelling those it could and censoring those it couldn’t cancel. The venality, dishonesty, and sheer lust for power and control that marked the Covid response – together with a *deeply* unimpressive record of actually getting the science right themselves – suggests that our science authorities are not to be trusted with policing science fraud, particularly as they’re often purveyors of fraudulent science themselves.

Yet there really is a problem. Science currently faces a [“replication crisis.”](#) in which vast numbers of published results don’t hold up when examined. Whole disciplines (\*cough\* social psychology \*cough\*) are so riddled with fraud as to be useless. And the public’s faith in science, which the “fraudbusters” of ORI were trying to preserve, has taken an enormous hit as a result.

Well, every successful system accumulates parasites, and American science has been enormously successful. But now it has accumulated a parasitic load that is rendering it weak and sick. So what do we do?

This being 2024, I asked an AI, Anthropic’s [Claude](#), for some proposed solutions to the replication crisis and got this. (My prompt: “What are some suggested solutions to the ‘replication crisis’ in science?”)

*The replication crisis in science is a significant issue that has gained attention in recent years. Here are some suggested solutions to address this problem:*

- 1. Improved statistical practices:*

- *Encouraging larger sample sizes*
- *Using more robust statistical methods*
- *Properly reporting effect sizes and confidence intervals*

## *2. Pre-registration of studies:*

- *Researchers outline their hypotheses, methods, and analyses before conducting the study*
- *Reduces the likelihood of p-hacking and HARKing (Hypothesizing After Results are Known)*

## *3. Open data and materials:*

- *Making raw data, analysis code, and materials publicly available*
- *Allows other researchers to verify and build upon existing work*

## *4. Replication studies:*

- *Encouraging and valuing direct replications of important findings*
- *Creating dedicated journals or sections for publishing replication attempts*

## *5. Changes in publication practices:*

- *Addressing publication bias by encouraging the publication of null results*
- *Implementing Registered Reports, where journals accept articles based on the methodology before results are known*

## *6. Improved research methods training:*

- *Enhancing education on research design, statistics, and scientific integrity*
- *Promoting best practices in graduate programs and continuing education*

## *7. Collaboration and multi-lab studies:*

- *Encouraging large-scale collaborations to increase sample sizes and diversity*
- *Conducting multi-lab replication projects to assess the generalizability of findings*

#### *8. Incentive structure reforms:*

- *Rewarding robust, reproducible research rather than just novel, positive findings*
- *Considering replication and open science practices in hiring, promotion, and funding decisions*

#### *9. Meta-analysis and systematic reviews:*

- *Emphasizing the importance of synthesizing multiple studies rather than relying on single experiments*

#### *10. Improved peer review:*

- *Training reviewers to focus on methodological rigor rather than just results*
- *Implementing open peer review to increase transparency in the publication process*

*These solutions aim to improve the reliability and credibility of scientific research across various disciplines. Implementing a combination of these approaches could help address the replication crisis and strengthen the scientific process overall.*

Some of these are not very specific – changes in publication practices? What would that mean? – and many boil down to just doing a better job, which would be nice, but seems hard to rely on. The most useful ones, in my opinion, revolve around changing incentives. Right now people get ahead by publishing, and people get published well, and publicized after publication, by achieving – or at least claiming – novel and interesting results. This produces an incentive to puff, to torture the data statistically so that it at least appears to indicate something, and even to just fake stuff, as the

Alzheimer's researchers reportedly [did](#) in boosting the amyloid plaque theory.

Asked to expand on the incentive structure reforms, [Claude](#) returned this (yes, I'm currently experimenting with AI-assisted research):

*Incentive structure reforms are a crucial aspect of addressing the replication crisis, as they aim to realign the rewards and motivations within the scientific community to promote more robust and reproducible research. Here's a more detailed look at this approach:*

### *1. Redefining success metrics:*

- *Shift focus from quantity to quality of publications*
- *Evaluate researchers based on the reproducibility and long-term impact of their work, not just the number of publications or journal impact factors*
- *Implement metrics that value open science practices, data sharing, and successful replications*

### *2. Funding allocation:*

- *Allocate specific funding for replication studies*
- *Require grant proposals to include plans for ensuring reproducibility*
- *Reward researchers who consistently produce reproducible results with increased funding opportunities*
- *Consider a researcher's track record in open science and replication when evaluating grant applications*

### *3. Hiring and promotion criteria:*

- *Include reproducibility and open science practices in faculty hiring decisions*
- *Consider a candidate's efforts in data sharing, pre-registration, and replication studies during tenure and promotion reviews*
- *Value methodological innovations that improve*

*reproducibility as much as novel findings*

4. *Publication practices:*

- *Encourage journals to adopt badges for open data, materials, and pre-registration*
- *Implement a system where successful replications can be linked to original studies, boosting the credibility and impact of reproducible work*
- *Create special issues or sections dedicated to replication studies and meta-analyses*

5. *Academic recognition:*

- *Establish awards and honors for researchers who consistently engage in open science practices*
- *Recognize reviewers who focus on methodological rigor rather than just novelty or positive results*
- *Create prestigious positions or titles for researchers who contribute significantly to improving research practices in their field*

6. *Collaborative incentives:*

- *Reward participation in large-scale, multi-lab studies*
- *Recognize contributions to shared resources, such as openly available datasets or analysis tools*
- *Encourage and value mentorship that focuses on robust research methods*

7. *Institutional policies:*

- *Implement institutional guidelines that require or strongly encourage open science practices*
- *Establish research integrity offices that provide support for reproducible research methods*
- *Create institutional repositories for data sharing and preprint publishing*

8. *Industry and public sector collaboration:*

- *Encourage partnerships between academia, industry, and government that prioritize reproducible research*
- *Implement standards for reproducibility in industry-sponsored research*
- *Tie public funding to adherence to open science practices*

#### *9. Education and training incentives:*

- *Offer grants or course reductions for faculty who develop curricula focused on robust research methods and reproducibility*
- *Provide recognition or certification for students who complete advanced training in reproducible research practices*

#### *10. Media and public engagement:*

- *Encourage researchers to communicate the importance of reproducibility to the public*
- *Reward clear reporting of scientific uncertainty and limitations in public-facing communications*

*By reforming these incentive structures, the scientific community can create an environment where rigorous, reproducible research is valued as much as, if not more than, novel and flashy results. This shift can help address the root causes of the replication crisis by aligning researchers' career incentives with practices that promote scientific integrity and reliability.*

Again, a mixture of the obvious, the trivial, and the useful. I like requiring researchers to specify ways of ensuring reproducibility in their applications, and evaluating researchers based on long-term reproducibility. As for "research integrity offices," well, that didn't work out so well the first time. Requiring data-sharing – and data "archiving," as it's surprising how often data for crucial studies turns out of have been lost in a move or a flood when

requested later – would also help.

And – and this was suggested by a commenter to an earlier blog post – not relying on scientific research for public policy purposes until it has been successfully replicated by someone else is not a bad idea. That would slow down the connection between research and public policy, but would that really be such a bad thing?

On the upside, there does seem to be something of a growing industry in trying to replicate studies, and in publishing the results – that’s one of the things that has led to the term “replication crisis.” And it’s also true that not all fields are equally afflicted: There’s much more of an issue in the biomedical fields and in psychology than in some other fields, like physics. I suspect that’s a product of both different incentives and different standards for research. But those are important fields, and likely influence public policy more than physics, too.

Well, having consulted AI, now I’m going to crowdsource this by consulting my readers. What do *you* think we should do to make scientific research more reliable?

*First published in [Glenn’s Substack](#)*