

# THE NEW YORKER

- « Video: Therapy on Chinese Television
- Main
- To-Do List: The Transition Edition »

January 3, 2011

## More Thoughts on the Decline Effect

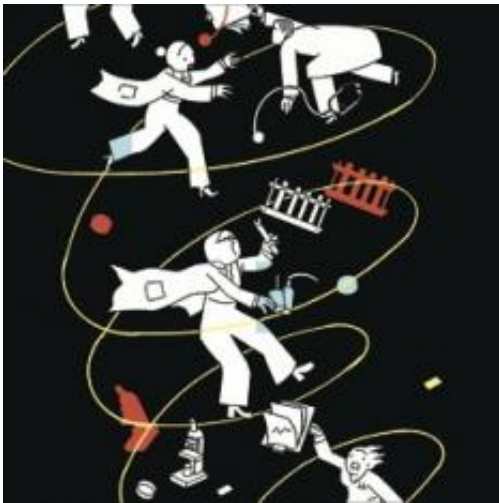
Posted by *Jonah Lehrer*

- Print
- More

### Share

Close

- 
- [Reddit](#)
- [Linked In](#)
- [Email](#)



In “The Truth Wears Off,” I wanted to explore the human side of the scientific enterprise. My focus was on a troubling phenomenon often referred to as the “decline effect,” which is the tendency of many exciting scientific results to fade over time. This empirical hiccup afflicts fields from pharmacology to evolutionary biology to social psychology. There is no simple explanation for the decline effect, but the article explores several possibilities, from the publication biases of peer-reviewed journals to the “selective reporting” of scientists who sift through data.

This week, the magazine published four very thoughtful letters in response to the piece. The first letter, like many of the e-mails, tweets, and comments I’ve received directly, argues that the decline effect is ultimately a minor worry, since “in the long run, science prevails over human bias.” The letter, from Howard Stuart, cites the famous 1909

oil-drop experiment performed by Robert Millikan and Harvey Fletcher, which sought to measure the charge of the electron. It’s a fascinating experimental tale, as subsequent measurements gradually corrected the data, steadily nudging the charge upwards. In his 1974 commencement address at Caltech, Richard Feynman described why the initial measurement was off, and why it took so long to fix:

Millikan measured the charge on an electron by an experiment with falling oil drops, and got an answer which we now know not to be quite right. It’s a little bit off, because he had the incorrect value for the viscosity of air. It’s interesting to look at the history of measurements of the charge of the electron, after Millikan. If you plot them as a function of time, you find that one is a little bigger than Millikan’s, and the next one’s a little bit bigger than that, and the next one’s a little bit bigger than that, until finally they settle down to a number which is higher.

Why didn’t they discover that the new number was higher right away? It’s a thing that scientists are

ashamed of—this history—because it's apparent that people did things like this: When they got a number that was too high above Millikan's, they thought something must be wrong—and they would look for and find a reason why something might be wrong. When they got a number closer to Millikan's value they didn't look so hard. And so they eliminated the numbers that were too far off, and did other things like that.

That's a pretty perfect example of selective reporting in science. One optimistic takeaway from the oil-drop experiment is that our errors get corrected, and that the truth will always win out. Like Mr. Stuart, this was the moral Feynman preferred, as he warned the Caltech undergrads to be rigorous scientists, because their lack of rigor would be quickly exposed by the scientific process. "Other experimenters will repeat your experiment and find out whether you were wrong or right," Feynman said. "Nature's phenomena will agree or they'll disagree with your theory."

But that's not always the case. For one thing, a third of scientific papers never get cited, let alone repeated, which means that many errors are never exposed. But even those theories that do get replicated are shadowed by uncertainty. After all, one of the more disturbing aspects of the decline effect is that many results we now believe to be false have been replicated numerous times. To take but one example I cited in the article: After fluctuating asymmetry, a widely publicized theory in evolutionary biology, was proposed in the early nineteen-nineties, nine of the first ten independent tests confirmed the theory. In fact, it took several years before an overwhelming majority of published papers began rejecting it. This raises the obvious problem: If false results can get replicated, then how do we demarcate science from pseudoscience? And how can we be sure that anything—even a multiply confirmed finding—is true?

These questions have no easy answers. However, I think the decline effect is an important reminder that we shouldn't simply reassure ourselves with platitudes about the rigors of replication or the inevitable corrections of peer review. Although we often pretend that experiments settle the truth for us—that we are mere passive observers, dutifully recording the facts—the reality of science is a lot messier. It is an intensely human process, shaped by all of our usual talents, tendencies, and flaws.

Many letters chastised me for critiquing science in such a public venue. Here's an example, from Dr. Robert Johnson of Wayne State Medical School:

Creationism and skepticism of climate change are popularly-held opinions; Lehrer's closing words play into the hands of those who want to deny evolution, global warming, and other realities. I fear that those who wish to persuade Americans that science is just one more pressure group, and that the scientific method is a matter of opinion, will be eager to use his conclusion to advance their cause.

This was a concern I wrestled with while writing the piece. One of the sad ironies of scientific denialism is that we tend to be skeptical of precisely the wrong kind of scientific claims. Natural selection and climate change have been verified in thousands of different ways by thousands of different scientists working in many different fields. (This doesn't mean, of course, that such theories won't change or get modified—the strength of science is that nothing is settled.) Instead of wasting public debate on solid theories, I wish we'd spend more time considering the value of second-generation antipsychotics or the verity of the latest gene-association study.

Nevertheless, I think the institutions and mechanisms of the scientific process demand investigation, even if the inside view isn't flattering. We know science works. But can it work better? There is too much at stake to not ask that question. Furthermore, the public funds a vast majority of basic research—it deserves to know about any problems.

And this brings me to another category of letters, which proposed new ways of minimizing the decline effect. Some readers suggested reducing the acceptable level of p-values or starting a *Journal of Negative Results*. Andrew Gelman, a professor of statistics at Columbia University, proposed the use of "retrospective power analyses," in which experimenters are forced to calculate their effect size using "real prior information," and not just the data distilled from their small sample size.

I also received an intriguing e-mail from a former academic scientist now working for a large biotech company:

When I worked in a university lab, we'd find all sorts of ways to get a significant result. We'd adjust the sample size after the fact, perhaps because some of the mice were outliers or maybe they were handled incorrectly, etc. This wasn't considered misconduct. It was just the way things were done. Of course, once these animals were thrown out [of the data] the effect of the intervention was publishable.

He goes on to say that standards are typically more rigorous in his corporate lab:

Here we have to be explicit, in advance, of how many mice we are going to use, and what effect we expect to find. We can't fudge the numbers after the experiment has been done... That's because companies don't want to begin an expensive clinical trial based on basic research that is fundamentally flawed or just a product of randomness.

Of course, once that basic research enters clinical trials, there's plenty of evidence that the massive financial incentives often start warping the data, leading to the suppression of negative results and the misinterpretation of positive ones. (This helps explain, at least in part, why such a large percentage of randomized clinical trials cannot be replicated. (ClinicalTrials.gov has tried to fix this problem by mandating the public registration of every clinical trial in advance.)

The larger point, though, is that there is nothing inherently mysterious about why the scientific process occasionally fails or the decline effect occurs. As Jonathan Schooler, one of the scientists featured in the article told me, "I'm convinced that we can use the tools of science to figure this"—the decline effect—"out. First, though, we have to admit that we've got a problem."

*Illustration: Laurent Cilluffo*

- Print
- More

## Share

Close

- 
- Reddit
- Linked In
- Email



## News Desk

- All posts
- RSS

11 hours ago

### Koch Pledge Tied to Congressional Climate Inaction

2 days ago

### A Year of Drawing Morsi

2 days ago

### Backlogged: Our Five Favorite Sentences of the Week

3 days ago

### The Crisis in a Nutshell: Quotes of the Week

- @tnynewsdesk
- Follow
- Our Newsletters

You might like

Most shared



Amazing, Previously Unpublished Mark Twain Quotes  
 BY JOE VEIX



G.O.P. Extremists Defy Description  
 BY JOHN CASSIDY



Poll: Americans Divided Over What Wild Animal They Would Like To See Congress Mauled By  
 BY ANDY BOROWITZ



Measuring America's Decline, in Three Charts  
 BY JOHN CASSIDY

To get more of *The New Yorker's* signature mix of politics, culture and the arts: **Subscribe Now**