

Reconsidering Science's 'Replication Crisis'



By *Chuck Dinerstein* — May 15, 2019



Courtesy of Daniel Sone, National Cancer Institute [1]

Science is an imperfect pursuit at its core. It advances in stops and starts, and only with time and trials do beliefs morph into models and then facts. That strength has been treated as a weakness by the other sources of authority, religion, and government, and at times used to cry fakery or chicanery. Science too, in more recent moments of self-reflection, has raised concerns about improperly fitting data to the hypothesis through hacking p-values and the lack of reproducibility in studies. In a fractal moment, a group of modelers created a model of scientific discovery seeking clues to what drives insight and innovation.

The group made use of computer simulations to consider a variety of factors that might influence our quest for factual truths. It is a sandbox model, of limited size and scope making it tractable to analyze, and like our own play in the sandboxes of our youth, it guides us towards and teaches necessary skills to build upon, it does not seek to be definitive. It is “an abstraction of reality,” containing “the salient features of the scientific process of interest to our questions.” The question they asked of their model was how quickly it arrived at the truth while varying

- The strength of the truth’s signal – how much noise was present
- How different scientists search for the truth – did they replicate prior work, did they alter factors in the currently accepted model or look for factors limited the model’s explanatory value or did they simply ignore the present beliefs and look elsewhere
- The complexity of the model – the number of variables required by the model
- The amount of work was replicated

The first finding should come as no surprise, the greater the signal, the more quickly the simulation identified the “truth.” How long we continued to hold a belief, the “stickiness” of that consensus, is

also high in the presence of a strong signal and generally, when the model is more complex, taking into account more variables. But that finding comes with some caveats. When more scientists apply a strategy looking for limitations, they spend more time creating unduly complex models that add no more significant information, and it takes longer to reach consensus. Every research strategy had flaws, and each spent varying amounts of time achieving consensus. The most effective search strategy was a balanced combination of all approaches.

This approach to scientific discovery achieved consensus when the search was not biased or strongly influenced by the current consensus, which I would argue is a fair description of our current approach to funding. We are more financially supportive of scientists exploring the limitations or nuances of our current beliefs, than of outlying theories. No one rushed to fund studies seeking an "infectious" cause for stomach ulcers; we deeply invested in studies of acid production. We know now that the most common cause is an infection by *Helicobacter pylori*.

It also requires that our measurement identify more signal than noise. As the authors point out this can be difficult in fields where data and signal are not tightly coupled, like psychology or where the instruments of measurement are insensitive; a problem in medicine as our criteria for disease alters with our instrumentation. [1]

When reproducibility was added to the researcher's simulation, it did not increase arriving at or maintaining consensus, it lent credibility but was insufficient to confirm a true model. And while it makes common sense that more time spent in replication reduces the speed in finding the true model, after all, there are only so many hours in the day, the role of replication is more nuanced because of the interactive effect of the search strategy. In an approach dominated by ignoring the consensus and searching elsewhere, there was little replication, and the truth was found quickly; when the boundary seekers were dominant, replication was much higher, but the truth was not as quickly ascertained. Replication may well be necessary but is not sufficient.

“Our study shows that even in this idealized framework, the link between reproducibility and the convergence to a scientific truth is not straightforward. ... both reproducibility and convergence to a scientific truth are presumably desirable properties of scientific discovery, they are not equivalent concepts. In our system inequivalence of these concepts is explained by a combination of research strategies, statistical methods, noise-to-signal ratios, and the complexity of the truth. This finding further indicates that issues regarding reproducibility or validity of scientific results should not be reduced down to questionable research practices or structural incentives.”

The search for scientific truth is like the truth it seeks, nuanced, interactive, a human activity. Replication has a role, but so does a diversity of theories and strategies for getting nature to give up its secrets.

[1] The recent spike in patients with “damaged” heart muscle reflects an increased sensitivity in testing rather than a clinical change.

Source: Scientific discovery in a model-centric framework: Reproducibility, innovation, and epistemic diversity PlosOne DOI: 10.1371/journal.pone.0216125

COPYRIGHT © 1978-2016 BY THE AMERICAN COUNCIL ON SCIENCE AND HEALTH

Source URL: <https://www.acsh.org/news/2019/05/15/reconsidering-sciences-replication-crisis-14026>

Links

[1] <https://visuals.nci.nih.gov/details.cfm?imageid=9427>