

# Beyond the 'replication crisis,' does research face an 'inference crisis'?

October 11 2019

---



Credit: CC0 Public Domain

For the past decade, social scientists have been unpacking a "replication crisis" that has revealed how findings of an alarming number of scientific studies are difficult or impossible to repeat. Efforts are

underway to improve the reliability of findings, but cognitive psychology researchers at the University of Massachusetts Amherst say that not enough attention has been paid to the validity of theoretical inferences made from research findings.

Using an example from their own field of memory research, they designed a test for the accuracy of theoretical conclusions made by researchers. The study was spearheaded by associate professor Jeffrey Starns, professor Caren Rotello, and doctoral student Andrea Cataldo, who has now completed her Ph.D. They shared authorship with 27 teams or individual cognitive psychology researchers who volunteered to submit their expert research conclusions for data sets sent to them by the UMass researchers.

"Our results reveal substantial variability in experts' judgments on the very same data," the authors state, suggesting a serious inference problem. Details are newly released in the journal *Advancing Methods and Practices in Psychological Science*.

Starns says that objectively testing whether scientists can make valid theoretical inferences by analyzing data is just as important as making sure they are working with replicable data patterns. "We want to ensure that we are doing good science. If we want people to be able to trust our conclusions, then we have an obligation to earn that trust by showing that we can make the right conclusions in a public test."

For this work, the researchers first conducted an online study testing recognition memory for words, "a very standard task" in which people decide whether or not they saw a word on a previous list. The researchers manipulated memory strength by presenting items once, twice, or three times and they manipulated bias—the overall willingness to say things are remembered—by instructing participants to be extra careful to avoid certain types of errors, such as failing to identify a

previously studied item.

Starns and colleagues were interested in one tricky interpretation problem that arises in many recognition studies, that is, the need to correct for differences in bias when comparing memory performance across populations or conditions. Unfortunately, this situation can arise if memory for the population of interest is equal to, better than, or worse than controls. Recognition researchers use a number of analysis tools to distinguish these possibilities, some of which have been around since the 1950's.

To determine if researchers can use these tools to accurately distinguish memory and bias, the UMass researchers created seven two-condition data sets and sent them to contributors without labels, asking them to indicate whether or not the conditions were from the same or different levels of the memory strength or response bias manipulations. Rotello explains, "These are the same sort of data they'd be confronted with in an experiment in their own labs, but in this case we knew the answers. We asked, 'did we vary memory strength, response bias, both or neither?'"

The volunteer cognitive psychology researchers could use any analyses they thought were appropriate, Starns adds, and "some applied multiple techniques, or very complex, cutting-edge techniques. We wanted to see if they could make accurate inferences and whether they could accurately gauge uncertainty. Could they say, 'I think there's a 20 percent chance that you only manipulated memory in this experiment,' for example."

Starns, Rotello and Cataldo were mainly interested in the reported probability that memory strength was manipulated between the two conditions. What they found was "enormous variability between researchers in what they inferred from the same sets of data," Starns

says. "For most data sets, the answers ranged from 0 to 100 percent across the 27 responders," he adds, "that was the most shocking."

Rotello reports that about one-third of responders "seemed to be doing OK," one-third did a bit better than pure guessing, and one-third "made misleading conclusions." She adds, "Our jaws dropped when we saw that. How is it that researchers who have used these tools for years could come to completely different conclusions about what's going on?"

Starns notes, "Some people made a lot more incorrect calls than they should have. Some incorrect conclusions are unavoidable with noisy data, but they made those incorrect inferences with way too much confidence. But some groups did as well as can be expected. That was somewhat encouraging."

In the end, the UMass Amherst researchers "had a big reveal party" and gave participants the option of removing their responses or removing their names from the paper, but none did. Rotello comments, "I am so impressed that they were willing to put everything on the line, even though the results were not that good in some cases." She and colleagues note that this shows a strong commitment to improving research quality among their peers.

Rotello adds, "The message here is not that [memory](#) researchers are bad, but that this general tool can assess the quality of our inferences in any field. It requires teamwork and openness. It's tremendously brave what these scientists did, to be publicly wrong. I'm sure it was humbling for many, but if we're not willing to be wrong we're not good scientists." Further, "We'd be stunned if the inference problems that we observed are unique. We assume that other disciplines and research areas are at risk for this problem."

**More information:** Jeffrey J. Starns et al, Assessing Theoretical

Conclusions With Blinded Inference to Investigate a Potential Inference Crisis, *Advances in Methods and Practices in Psychological Science* (2019). [DOI: 10.1177/2515245919869583](https://doi.org/10.1177/2515245919869583)

Provided by University of Massachusetts Amherst

Citation: Beyond the 'replication crisis,' does research face an 'inference crisis'? (2019, October 11) retrieved 20 April 2024 from <https://medicalxpress.com/news/2019-10-replication-crisis-inference.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.