

June 23, 2014

Replication Crisis in Psychology Research Turns Ugly and Odd (Chronicle of Higher Education)



By Tom Bartlett

Another salvo was fired recently in what's become known in the psychology Twitterverse as "repligate."

In a blog post published last week, Timothy D. Wilson, a professor of psychology at the University of Virginia and the author of [Redirect: The Surprising New Science of Psychological Change, declared](#) that "the field has become preoccupied with prevention and error detection—negative psychology—at the expense of exploration and discovery."

The evidence that psychology is beset with false positives is weak, according to Mr. Wilson, and he pointed instead to the danger of inept replications that serve only to damage "the reputation of the original researcher and the progression of science." While he called for finding common ground, Mr. Wilson pretty firmly sided with those who fear that psychology's growing replication movement, which aims to challenge what some critics see as a tsunami of suspicious science, is more destructive than corrective.

Still, Mr. Wilson was polite. Daniel Gilbert, less so. Mr. Gilbert, a professor of psychology at Harvard University, author of the best-seller *Stumbling on Happiness*, and host of the PBS show *This Emotional Life*, [wrote](#) that certain so-called replicators are "shameless little bullies" and "second stringers" who engage in tactics "out of Senator Joe McCarthy's playbook" (he later took back the word "little," writing that he didn't know the size of the researchers involved).

Scrutiny From the Replicators

What got Mr. Gilbert so incensed was the treatment of Simone Schnall, a senior lecturer at the University of Cambridge, whose [2008 paper on cleanliness and morality](#) was selected for replication in a special issue of the journal *Social Psychology*.

The Schnall case has been dissected in [multiple venues](#), but it won't hurt to replicate it here. In one experiment, Ms. Schnall had 40 undergraduates unscramble some words. One group unscrambled words that suggested cleanliness (pure, immaculate, pristine), while the other group unscrambled neutral words. They were then presented with a number of moral dilemmas, like whether it's cool to eat your dog after it gets run over by a car. Ms. Schnall wanted to discover whether prompting—or priming, in psych parlance—people with the concept of cleanliness would make them less judgmental.

Turns out, it did. Subjects who had unscrambled clean words weren't as harsh on the guy who chows down on his chow. This is consistent with [other research by Ms. Schnall](#) that found that subjects were more judgmental when they were in a dirty room or had to smell fart spray (which, unfortunately, is [a real product](#)).

These studies fit into a relatively new field known as embodied cognition, which examines how one's environment and body affect one's feelings and thoughts. Previous studies have found, for instance, that holding a heavy clipboard makes people feel that the task at hand is weightier and that being socially excluded increases a desire for warm foods. Part of the idea is that the metaphors we use, such as associating weight with importance or friendliness with warmth, are grounded in measurable psychological reality.

Perhaps because the phenomenon seems odd, priming and embodied cognition have received special scrutiny from the replicators. Michigan State University researchers [reran](#) Ms. Schnall's experiment, mostly mirroring her procedures but increasing the number of subjects from 40 to 208. The effect Ms. Schnall had detected disappeared. The authors wrote that "we found little support for the idea that cleansing behaviors impact moral judgments." In short, the study failed to replicate.

Selective Science?

Ms. Schnall [objected at length](#) to the replication of her study. Her central complaint was that Michigan State researchers' replication was hampered by the "ceiling effect," which is to say that, according to her, too many of their subjects gave extreme responses to the moral dilemmas and therefore obscured any differences between the experimental group and the control. So the replicators ran the numbers again, eliminating the extreme responses, but they still didn't see an effect. Ms. Schnall objected to that effort as well, arguing that the researchers had thrown out a huge chunk of their data and therefore had engaged in the kind of selective science that they condemned.

One of the Michigan State researchers, Brent Donnellan, [pointed out in a blog post](#) that, in at least one instance, the supposedly extreme responses they received were the same as Ms. Schnall's. "I worry that any ceiling argument only applies when the results are counter to the original predictions," he wrote. Now, when you think about it, this dispute seems healthy enough. Lots of back and forth. Researchers double-checking one another's work. Journal editors even went so far as to publish a [long series of email exchanges](#) between the parties involved. If only all science was this interactive and transparent! But Ms. Schnall also accused the replicators of bullying. She had been, she wrote, inundated with requests for her data from replicators who were interested only because they suspected something might be wrong. "A truly [sic] scientific approach would be to randomly sample from the entire field, and conduct replications, rather than focus on the same topics (and therefore the same researchers) again and again," she wrote.

Ms. Schnall also complained about [a blog post](#) written by Mr. Donnellan that summarized his and his fellow researchers' findings. The post was titled "Go Big or Go Home," a reference to their contention that Ms. Schnall's sample size was too small and therefore unreliable. He wrote the following in the post: "We gave it our best shot and pretty much encountered an epic fail as my 10-year-old would say."

Calling Out the Bullies

This is surely what Ms. Schnall was referring to when she wrote that it is "uncollegial, for example, to tweet and blog about people's research in a mocking tone, accuse them of questionable research practices, or worse. Such behavior amounts to bullying, and needs to stop." Daniel Gilbert rode to Ms. Schnall's defense, agreeing that this was bullying and accusing Mr. Donnellan and others of "advertising their pleasure in in other's [sic] misfortunes." In a comment on Ms. Schnall's post, Mr. Gilbert likened Ms. Schnall to Rosa Parks

because she, too, was "a powerless woman who has decided to risk everything to call out the bullies."

Leaving aside the aptness of that comparison, Ms. Schnall is not, as Mr. Gilbert wrote originally, a young untenured professor (a mistake he corrected): She is a senior lecturer with tenure at one of the world's most prestigious universities. Which doesn't mean she can't be bullied, just that she's not a junior academic hunting for steady work.

Ms. Schnall told me, via email, that the replicators had never adequately addressed the issue of extreme responses. She sent me excerpts from some of the notes of support she had received from other researchers who themselves were the targets of what they considered biased replications. "It always seems very unfair that the attackers go completely free if they accuse in error," one researcher wrote to her.

One outspoken replicator, Harold Pashler, a professor of psychology at the University of California at San Diego, pushed back against the allegation from Mr. Wilson, Mr. Gilbert, and others that replicators are—to use Mr. Wilson's term—"ill intentioned." "He should tell us on what basis he uses that phrase," Mr. Pashler wrote in an email. "Everyone I know who's been trying to replicate social-cognition results is genuinely curious about the results and is quite scrupulous in executing these designs with care."

A Peace Broker

It seems that social psychology, and perhaps science in general, is in the midst of a shift toward greater accountability. That's a good thing. But it's also largely uncharted territory. Daniel Kahneman, a Nobel Prize winner who has tried to serve as a sort of a peace broker, recently offered [somerules of the road](#) for replications, including keeping a record of the correspondence between the original researcher and the replicator, as was done in the Schnall case. Mr. Kahneman argues that such a procedure is important because there is "a lot of passion and a lot of ego in scientists' lives, reputations matter, and feelings are easily bruised."

That's undoubtedly true, and taking glee in someone else's apparent misstep is unseemly. Yet no amount of politeness is going to soften the revelation that a published, publicized finding is bogus. Feelings may very well get bruised, reputations tarnished, careers trashed. That's a shame, but while being nice is important, so is being right.

-