

Revisiting Stereotype Threat - by Michael Inzlicht

Author: Michael Inzlicht

Date: Dec 18, 2024

Source:

<https://www.speakandregret.michaelinzlicht.com/p/revisiting-stereotype-threat>

Another day, another idol falls.

This one has been teetering for years, so the collapse didn't come as a shock. But that doesn't make it any less painful.

I'm talking about stereotype threat, a once-revolutionary idea that shaped how social psychologists thought about identity, achievement, and inequality. For decades, it inspired research, drove interventions, and promised insights into the invisible forces that constrain human potential.

I still remember seeing its most eloquent advocate, Stanford University's Claude Steele, deliver a keynote address in 1999 at the annual convention of what was then called the American Psychological Society. It was my first ever conference, my first trip to Denver, and Steele was nothing short of magnetic. Charismatic and at the height of his powers, he commanded the stage like no academic I had ever seen. He delivered his message with the kind of confidence that makes you believe science can change the world. Professor Steele was a rock star, and I was as giddy seeing him on stage as I was seeing Kurt Cobain on stage a few years earlier.

What is Stereotype Threat?

The concept of stereotype threat, first proposed by Claude Steele in the early 1990s, posited that individuals who are part of a negatively stereotyped group can, in certain situations, experience anxiety about confirming those stereotypes, leading paradoxically to underperformance, thus confirming the disparaging stereotype. The initial research was groundbreaking.

In 1995, Steele and his student Joshua Aronson—who went on to become my postdoc supervisor years later—demonstrated that the notorious Black-white gap in academic performance could be partially closed when negative stereotypes impugning Black people's intelligence were made irrelevant. When Black students at Stanford University were told that a test was diagnostic of intellectual ability, they performed worse than their white counterparts.

However, when this *stereotype threat* was ostensibly removed—by simply framing the test as a measure of problem-solving rather than intelligence—the performance gap Black and white students nearly vanished.

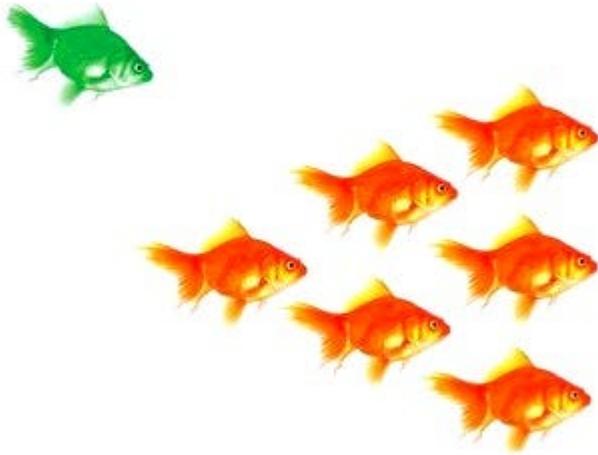
Share

Suddenly, here was an explanation for why certain groups didn't perform as well in academic settings. And it wasn't just race; follow-up studies looked at women in math and science. Women, who dominate men in most academic disciplines, underperform in STEM fields because they were regularly, albeit subtly, reminded of the stereotype that women aren't good at math, or so the story goes. The idea felt revolutionary, hopeful even, because it suggested that these vexing performance gaps could be addressed by changing people's immediate environments rather than accepting them as fixed outcomes, inherent to the groups themselves

[

STEREOTYPE THREAT

THEORY, PROCESS, AND APPLICATION



EDITED BY
MICHAEL INZLICHT AND TONI SCHMADER

OXFORD

Books — Michael Inzlicht

]

(https://substackcdn.com/image/fetch/f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2Fd8afdc60-dc85-4132-a28d-7eeb4c378328_355x550.jpeg)

These findings were exhilarating. Before long, stereotype threat was not only the darling of social psychology, but it also became the darling of the political left who now had an answer to prevailing views of group differences held by the political right. This is partly because shortly before stereotype threat took its turn in the spotlight, Charles Murray and Richard Herrnstein published *The Bell Curve*, which resulted in a media firestorm that has had repercussions to this day. Not only did the book discuss racial differences in intelligence as real and consequential—and not mere products of culturally biased IQ tests—it suggested that a non-negligible factor in this gap was due to biological

differences. This thesis was so toxic that the octogenarian Murray is still considered a pariah, shouted down and deplatformed from talks he tries to deliver at respectable colleges to this day.

Stereotype threat, in contrast, was a breath of fresh air. It promised that group differences were malleable, not fixed. They could be explained as momentary apprehension, akin to the nerves that might cause an elite athlete to choke on competition day. Yes, these group differences still have consequences, but now we have a remedy—change the situation so that stereotypes are less likely to be in the air and watch as all the Black students and female mathematicians rise to the top.

I too was swept up by this mania. I studied stereotype threat as a PhD student and published some of the first papers on the topic. My dissertation and very first publication suggested that subtle aspects of a room—like how many men and women were in a math classroom—could be enough to evoke stereotype threat and undermine performance. Because my field became captivated by stereotype threat, this meant that I was quickly offered jobs, grants, tenure, and acclaim. I edited a book on stereotype threat and was asked to add my name and research to briefs delivered to the US Supreme Court. My career benefitted immensely.

The Replicators are Coming

Then things started going sideways. And not just for stereotype threat.

It's all very complicated. Lots of strands in old Duder's head. But here's the skinny. In the early 2010s, psychology started looking inward, asking hard questions about the robustness of our most cherished findings. This happened because in the early 2000s, our best journals regularly included studies that were ludicrous and hard to believe. For example, researchers made dubious claims about the role of blood glucose and self-control, and found positive evidence for these claims, despite their biological impossibility. A paper was published in social psychology's most prestigious journal claiming evidence for clairvoyance, essentially offering a ringer for a ringer. Ludicrous. If these impossible ideas were generating support with the standard methods of social psychology, maybe our methods are not what we thought they were. A small cadre of reformers then started raising awareness that all was not right in how we conducted our science: we did not bother replicating important studies, we were misusing and abusing our statistical tools, and we did not publish all our studies—particularly the failed ones. And when some brave scientists decided to audit the field by closely replicating many studies, only about a quarter from social psychology could be successfully replicated. Since these dark days, the field has changed immensely, and we're slowly producing more respectable science today.

Share Speak Now Regret Later

Nonetheless, the entire field's evidentiary basis was now suspect. After all, they were produced by methods that we now consider questionable. Stereotype threat was no different. I would love to say that stereotype threat was an exception, that it survived replication attempts and other audits, and that a beloved idea can still be used to counter damaging claims about group differences. But new data now reveal what many of us suspected for at least ten years: stereotype threat does not replicate, and it does not undermine academic performance in the ways we thought.

This new data emerged from what is called a Registered Replication Report. This was no ordinary replication study; it used the gold standard of scientific rigor. Conducted by multiple labs across the U.S. and Europe, and led by Andrea Stoevenbelt this study (still a preprint) was preregistered (meaning all methods and analyses were specified before the data were collected) and involved over 1,500 participants. It replicated the exact procedures of a well-known stereotype threat study published in 2005 by Mike Johns, Toni Schmader, and Andy Martens—all colleagues and friends I deeply respect. The original study had found that women performed worse on math tests when reminded of gender stereotypes but performed on par with men when they were instead taught about stereotype threat. The idea was that awareness of the phenomenon of stereotype threat helped mitigate its effects, which was why this original paper was so influential: it offered a simple intervention to close the gender-gap in math performance. The replication was designed to be thorough, with consistent methodology across sites and a sample size large enough to detect even small effects.

Despite following these procedures to the letter, the replication found no effect. Women who were ostensibly in a threat condition didn't perform any worse than those who were instead taught about threat. And the difference between men and women's math performance remained consistent across the board, regardless of how the test was framed. The stereotype threat effect, once thought to be so robust, just wasn't there.

What Does This Mean for Stereotype Threat?

Does one failed replication debunk the entire theory of stereotype threat? No, of course not. But it's not just one study. There are now multiple failed replications, large-sample studies that found no effect, and at least one bias-corrected meta-analysis pointing to the same conclusion: if stereotype threat exists, it is far weaker and more inconsistent than we originally believed. I no longer believe it is real, but you can make up your own mind.

I have seen some people online suggest the reason this failed to replicate is that women are no longer stereotyped as not being good at math. While I do not disagree that cultural stereotypes about women in STEM might have changed since 2005 when the original paper was first published, I'm skeptical this is the main culprit behind this non-replication. First, women remain heavily outnumbered in STEM fields. The latest statistics indicate that women comprise only 25% of STEM workers in Canada and 27% in the US. And, depending on what is counted as STEM—I have heard some argue that psychology should be included—this number might be a lot lower. So, the stereotype about what is and what is not a female job might still be around, as much as we'd like it not to be.

Second, for years, many of us have suspected that something wasn't right. There were warning signs: tiny sample sizes, flexible analyses, and implausibly large effect sizes given the relatively modest interventions being tested. In some cases, stereotype threat effects were found only in very specific handpicked samples—another red flag. It turns out that many of the original studies were conducted at a time when researchers—and I count myself here—were less stringent about methodological rigor.

Let's be honest: that last sentence was far too generous. Many of us engaged in practices that, in hindsight, were borderline dishonest. We abused experimenter degrees of freedom, engaged in questionable research practices, p-hacked, massaged our data—you pick the euphemism. In contrast, this new replication study followed the most up-to-date best practices in psychological science, eliminating room for flexibility in analysis or results interpretation.

In my opinion—one that I have shared widely over the years—studies like this do more than demonstrate that stereotype threat is not replicable. They raise unsettling questions about the broader field of social psychology. If stereotype threat is not real, not robust, what else was I taught in my introduction to psychology classes that is also suspect? Despite all our improvements that help us in the present and future, we still have a massive backlog of studies from the past that we need to reckon with. Yes, the future looks bright, but we need to have the courage to put our most cherished findings under the spotlight.

The bill of reckoning for social psychology is past due.

A Reckoning...and a Path Forward

The fall of stereotype threat is not just about one theory collapsing: it's a moment of reckoning for the entire field of social psychology. Stereotype threat was more than an idea—it was a promise, a way to understand inequality and to

imagine solutions. Its failure forces us to confront uncomfortable questions about how science is done and what happens when beloved ideas turn out to be wrong.

[Leave a comment](#)

But this reckoning, painful as it is, should not lead us to despair. The scientific process thrives on self-correction, on challenging old paradigms and building stronger ones in their place. What we're experiencing now is science doing what it's supposed to do: correcting itself. If we care about understanding the human mind and addressing real-world inequalities, we need to keep asking hard questions and demanding better evidence—not just for stereotype threat but for every cherished finding.

For me, letting go of stereotype threat has been both humbling and liberating. It has forced me to recalibrate how I think about research, advocacy, and the stories we tell about human potential. It's a reminder that science, at its best, is about progress, not protecting idols.

Thanks to everyone who's already shared their stories for *The Regret Project*—I've been blown away by how honest and moving some of the submissions have been. This project is a space to reflect on those moments we all have: regrets, missteps, and things we wish we could take back. Some of these stories will be featured in the next installment of *Regret Now*, so stay tuned. If you haven't shared yours yet, there's still time to contribute (anonymously!) [at this link](#). And when I say anonymously, that means I don't know your identity either. So no one, including me, will know your identity. So, let's take those moments of regret and turn them into something meaningful—or at least something that feels a little less isolating.

Let's play "Find the *Lebowski* quotes game" again! The first two Urban Achievers who spot the two Easter eggs in this post will get three months of free paid subscription. You know the rules, one quote per person; no using ChatGPT. No need to roll on Shabbos—just be the first to comment below with the quotes you found. Let's see who really abides.