

OPINION

Psychology research skews mostly white and wealthy

BY MICHAEL SCHULSON, UNDARK

27 JANUARY 2020

When **Cristine Legare** gives talks to groups of psychology researchers, she likes to take a quick poll of the room. How many of them, she asks, consider themselves to be "Western ethnopsychologists?" The question does not go over well. "They're like, 'What?'" Legare, associate professor of psychology at the University of Texas at Austin, says. "It doesn't resonate at all."

That confusion is precisely Legare's point. For decades, the overwhelming majority of psychology research has examined people who live in the United States and other affluent Western countries. By focusing on such a narrow population, Legare and other critics argue, psychology researchers have — mostly unwittingly — presented a skewed view of the human mind.

"It's not that it's not interesting or useful to study your American middle-class population. But they don't want to claim they're *just* studying that population," Legare says. "They want to claim that humans are alike enough that it really doesn't matter which population you study." Often, psychology papers do not even mention the nationality or socioeconomic status of, or other basic demographic statistics about, the study participants.

In many cases, other research suggests, the population being studied does matter — often in **subtle and profound ways** — and Legare is not the first researcher to voice these concerns. Debates about the diversity of psychology study participants reached a peak around 2010, when a **widely read paper** charged that an overreliance on research from Western, educated, industrialized, rich and democratic societies — often shortened to the acronym 'WEIRD' — amounted to a crisis for the behavioral sciences. At the time, it seemed possible that the field would undergo major reforms.

A decade later, however, many psychologists say that little has changed. In the process, they are raising questions about how, exactly, psychology researchers should account for nationality, class, gender, sexuality, race and other identities in their work — and expressing frustration at the lack of concrete reform.

“It’s the issue that we all like to talk about,” says University of Kentucky psychologist **Will Gervais**, “but nobody likes to actually change.”

Sounding alarms:

Beginning in the early 20th century, psychology researchers — who, in the field’s first decades, had mostly experimented on themselves — began to seek larger samples. In many cases, they turned to the most convenient captive populations they had at hand: local schoolchildren or undergraduates at their own universities. Given that recruiting people to participate in studies can be both difficult and costly, such close-to-home recruiting has persisted to the present day, though it is now sometimes augmented by services like Mechanical Turk, or MTurk, an Amazon platform that connects freelance workers (read: study participants) with low-wage menial tasks.

Whatever the source, these samples, at least on university campuses, typically skew toward white and affluent populations. They also draw heavily from industrialized Western societies. And yet, researchers often downplay the social identities of participants in published research — a tack that serves to highlight the universality of their results. “It became customary to emphasize the experimental identity of human data sources at the expense of their ordinary personal and social identity,” **writes historian Kurt Danziger** in “Constructing the Subject,” a classic 1990 study.

Researchers had some good reasons to be hesitant about emphasizing identities such as race or nationality. There’s a long history of scientists seeking to bolster racist, xenophobic arguments by positing, without any actual evidence, deep-seated differences among groups. Especially after World War II, intellectual currents swung in the opposite direction, emphasizing the universality of human experience.

And often those other identities don’t particularly matter. “A lot of what we do is consistent across people,” says **Daniel Simons**, professor of psychology at the University of Illinois in Champaign, who has written about generalization in psychology.

Early psychology research, Simons points out, often focused on basic behaviors that were unlikely to be influenced by culture or environment. Over time, researchers began to study more complex, social behaviors and “continued to assume that it was the same sort of universal principles.” Today, a lot of psychology research does examine topics where culture or particular experiences might shape or inform results — indeed, culture and environment might be at the very center of the issue. And on a lot of questions, Simons says, “we just don’t know.”

Given that knowledge gap, some psychologists have been sounding alarms for years. In the late 1990s, psychologist **Stanley Sue** expressed concern that his field paid too little attention to the experiences of non-white ethnic groups. A 2008 study that found that the research in six major psychology journals only **rarely examined people outside the West**, wryly proposed that a top journal rename itself the "Journal of the Personality and Social Psychology of American Undergraduate Introductory Psychology Students."

WEIRD science:

The issue gained traction in 2010, when **Joseph Henrich** and two colleagues at the University of British Columbia in Canada marshaled evidence from dozens of studies to demonstrate that people who grew up in the so-called 'WEIRD' societies often behave quite differently from people in other parts of the world. For example, certain optical illusions that consistently fool people from industrialized countries simply do not trick people who grow up in rural, non-industrialized societies. Or when asked to play a game that involves sharing money with a stranger, American undergraduates act differently from members of the Tsimane people, who live in the Bolivian Amazon.

"If the database of the behavioral sciences consisted entirely of Tsimane subjects, researchers would likely be quite concerned about generalizability," Henrich and his colleagues wrote. Why, they wondered, were researchers less concerned when their databases were made up almost entirely of Americans and Europeans?

The paper generated countless responses, meetings and calls for reform. Widely covered in the media, it has since been cited thousands of times in the scholarly literature. But Henrich, now a professor at Harvard University, says that so far the paper has had little effect on psychology research as a discipline. "On one level, I feel like there's a lot more enthusiasm around addressing sample variability," Henrich says. "But if you actually look at the numbers, the latest numbers coming in the last few years don't actually show any shift in the diversity of samples."

Some research backs him up. One recent analysis of papers published in *Psychological Science* found that, of the studies that even noted the nationality of participants, 94 percent focused exclusively on WEIRD samples. And more than 90 percent did not offer any data about participants' socioeconomic status.

Over the past 10 years, psychology has undergone a seismic change — just not exactly the one Henrich and others envisioned. Researchers began to realize that when they redid many major studies in the field, they could not replicate the results. Shoddy experimental practices and bad statistical habits, which helped to make random fluctuations in the data seem like big, meaningful results, were largely blamed for this replication crisis. But another and less frequently surfaced culprit, some psychologists say, is the lack of diversity in the original research samples: Studies tested in one population were simply not working in other populations.

“In my mind, those two things have always gone together,” says **Neil Lewis, Jr.**, assistant professor of communication and social behavior at Cornell University, of the relationship between sample diversity and the replication crisis.

And yet, when psychologists launched major efforts to replicate old research and to reform their experimental practices, critics say, they paid less attention to the lack of diversity in their samples. “Figuring out whether or not the finding you have today really works in all places and with all populations has not really been incentivized,” Lewis says.

Instead, some reform-minded psychologists suggest, it can seem as if the field continues to favor quick, flashy research over conscientious improvements in study design. In many institutions, “the reward structure is such where I would get ahead by publishing 20 crappy MTurk studies instead of one big cross-cultural one,” says Gervais, the Kentucky psychologist. “I don’t think we’d learn 20 times as much, but my CV would look better.”

Publishing pressure:

There is also pressure, some researchers say, to draw large, universal lessons from studies. “We’re really encouraged to make these big bold claims, and to have what feels like these groundbreaking papers,” says **Jasmine DeJesus**, assistant professor of psychology at the University of North Carolina at Greensboro, who has documented the prevalence of **broad, unsupported claims** in psychology papers.

Adding to the challenge, since the replication crisis began, psychology researchers have been increasingly expected to use bigger samples sizes in their research. Those new standards have been widely praised for increasing rigor in the social sciences, but they can place additional burdens on researchers who study underrepresented populations, which can be more difficult and expensive to recruit.

Taken together, these challenges can be formidable. **Sarah Gaither**, assistant professor of psychology and neuroscience at Duke University in Durham, North Carolina, studies identity, including how people conceptualize racial categories. Much of her work focuses on biracial children, a group that is often shunted into one racial category, or simply excluded from studies altogether. But biracial children’s experiences, Gaither’s research suggests, subvert some assumptions about how people’s minds come to conceptualize difference.

“The second you look at non-white people, you find very different effects on how they see these multiracial faces,” says Gaither, who pointed out that most of the work on the psychology of racial categorization has taken place on predominantly white samples. “Without having a diverse sample, you would never know that, because the majority of our papers don’t even report race and ethnicity properly.”

Gaither says she entered the field of psychology because she wanted to study underrepresented groups. Lacking tenure, however, she says she feels the need to publish frequently, compelling her to spend more time on studies that focus on predominantly white samples, recruited online.

And even when studies on underrepresented groups are done, Gaither adds, they typically attract less attention. “If you do study an underrepresented group, you’re just naturally not going to have the kind of citation count as someone who’s studying a more mainstream question,” Gaither says. That’s because researchers who are otherwise quick to extrapolate from predominantly white samples, her experience suggests, may be less likely to do so when the group is more diverse. Instead, the study ends up in a specialty journal that focuses on minority groups, where it may get fewer citations. “If you’re not studying black people, there’s no reason why you would want to cite a paper looking at black participant outcomes, for example.”

Necessary squirming:

Some reform efforts are underway. The Psychological Science Accelerator is a new global effort that takes specific experimental findings and then tests them in dozens of cultural contexts around the world. As Dalmeet Singh Chawla reported in *Undark* in November, the effort recently released its first study, which used more than 11,000 participants from 41 countries to replicate an influential 2008 experiment on how humans **judge the faces of strangers**.

Other, more modest attempts to address the issue focus on reforms at the academic journals that publish scientists’ work. Simons, the Illinois psychologist, has suggested that psychology papers adopt an entirely new section for what he calls Constraints on Generality, or COG, statements, which require researchers to define exactly which populations their research applies to. Other psychologists have urged journals to set explicit policies that favor research that includes underrepresented and non-WEIRD groups — even, perhaps, setting quotas to ensure that research represents a broader slice of humanity.

Some leaders of the field’s most influential institutions have heard those criticisms. “We have to embrace the need for greater diversity in our samples,” says **Patricia Bauer**, professor of psychology at Emory University in Atlanta, Georgia, who this month started a four-year term as editor in chief of *Psychological Science*.

Still, Bauer stresses that changes will take time. She points to the recent calls for 50 percent of papers in the journal to involve non-WEIRD participants in 2020. “I don’t think I can reach that goal,” she told *Undark*. “I think that’s too high. But having that in my mind, that will cause me to take certain steps.”

Bauer, who had not yet assumed the editorship of the journal when interviewed by *Undark*, shared some thoughts on what those steps might be. They include appointing a more diverse editorial board; sending signals that research in non-WEIRD populations is important work, a theme that

Bauer brought up forcefully in her **first editor's note** in the journal; and, perhaps, pushing authors to do more to justify why they pick the groups they do. Proposals like mandatory COG statements or other fixed policies, though, give her pause: "I don't like requirements," she says.

Bauer stresses that researchers have to balance competing needs, citing some of her own recent research on educational outcomes in a community in the American South that is roughly one-third black, one-third Latino and one-third white. By lumping everyone together, Bauer has a large enough sample to do the sorts of analyses that allow researchers to identify meaningful results from statistical noise. But if she were to try to break down the population by race or socioeconomic status, each group would be too small to actually analyze.

"I sometimes tell my students, don't put information in there about your sample that's going to cause a reviewer to ask you to analyze your data by subgroups, because our studies are not [statistically] powered that way," Bauer says.

Those kinds of suggestions are unlikely to win support among people pushing for more attention to sample diversity, who might want more disclosure, or bigger sample sizes, rather than simply leaving out the information. "That's grim," Henrich said when I told him about Bauer's advice to students.

For some reform-minded psychologists, leaders in the field cannot respond quickly enough. Legare says there is still an unspoken assumption that the most legitimate studies — the ones that best point to universal truths — are those that use white, English-speaking participants.

"There's some really uncomfortable ethnocentrism associated with this issue that makes people squirm," Legare says. "We should all be doing a lot more squirming."

*This story originally appeared on **Undark**. It has been slightly modified to reflect Spectrum's style.*