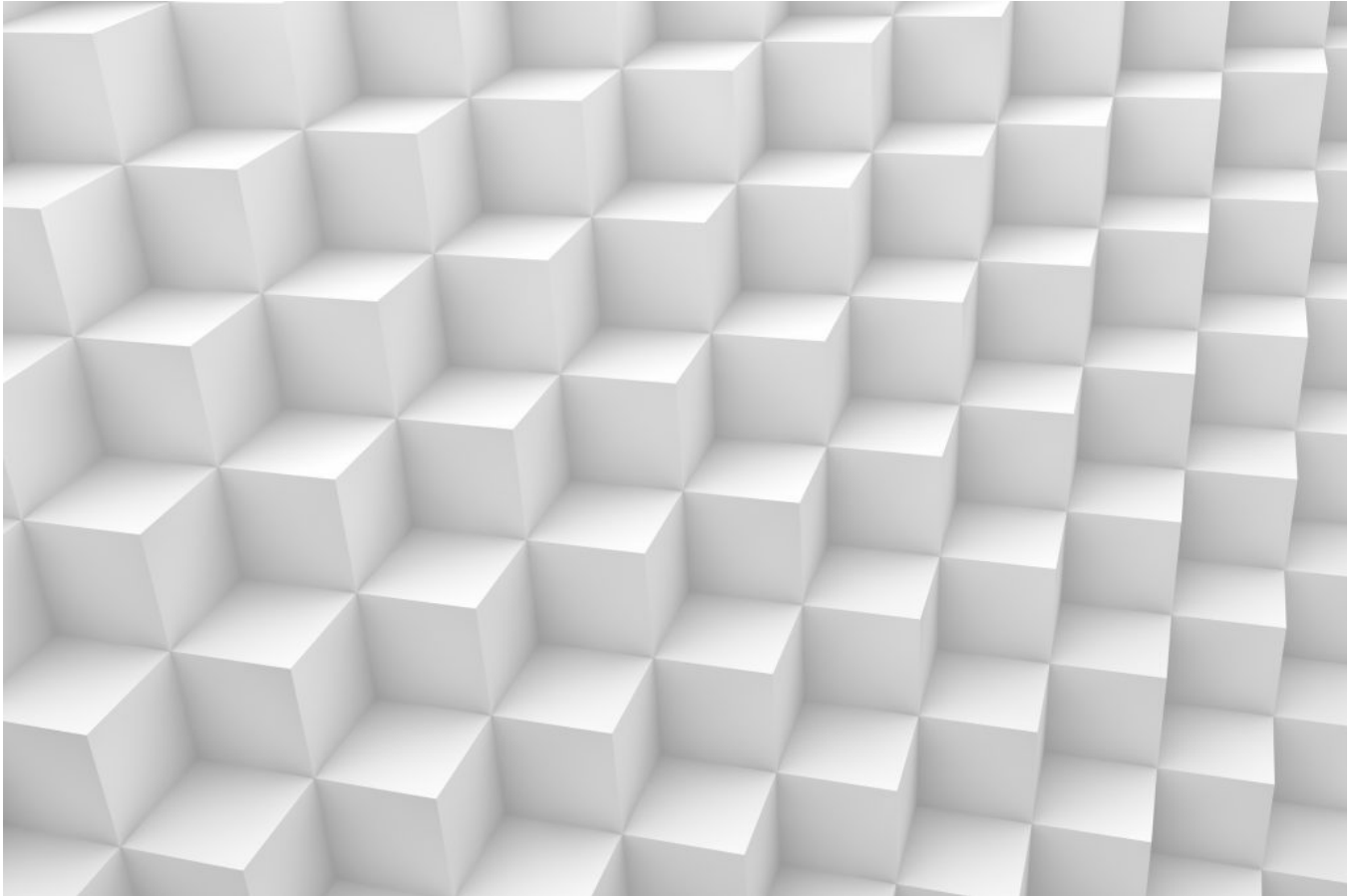


Replications of Psychological Findings Don't Appear to Hinge on Study-Population Differences, Global Multilab Project Shows

November 19, 2018



Failures to reproduce psychological research findings are often attributed to differences in the study population being examined. But results from a massive research project upend that claim.

A report on the international project, which involved replications of 28 classic and contemporary findings in psychological science, is forthcoming in *Advances in Methods and Practices in Psychological Science*, a journal of the Association for Psychological Science. A team of 186 researchers involved in the effort found that population characteristics had no bearing on the failure of a finding to replicate.

The project, called Many Labs 2, was designed specifically to address the argument that variations in study samples and research procedures may result in a failed replication. Each of the 28 studies was repeated in more than 60 labs across 36 nations and territories, and collectively generated sample sizes that on average were more than 60 times larger than the original samples.

Fourteen of the original findings replicated, although some at variable degrees across the different labs.

But for the 14 studies that did not replicate, sample diversity had minimal if any effect on the results.

“We were surprised that our diversity in our samples from around the world did not result in substantial diversity in the findings that we observed” said Rick Klein, a researcher at the University of Grenoble Alpes in France and one of the project leaders, in a public statement. “Among these studies at least, if the finding was replicable, we observed it in most samples, and the magnitude of the finding only varied somewhat depending on the sample.”

The studies selected included many studies published within the last 20 years but also some classics in the research literature, including the well-known framing effect on choices, identified by Amos Tversky and Daniel Kahneman; and a 1977 finding on the false consensus effect, in which people overestimate the consensus around their own beliefs and preferences. (Both those findings were reproduced, although the framing effect proved only half as robust in the replications compared to the original finding.)

Many Labs 2 represents a far greater undertaking compared to the first Many Labs project from 2013, in which 36 labs collaborated to examine 13 findings – replicating 10 of them. As part of the initiative, the collaborating labs collected the original materials from each study and had the experimental procedures peer-reviewed in advance by experts and, in some cases, authors on the original work.

The results of the latest project do not definitively mean the original findings were invalid, said Michelangelo Vianello, a professor at the University of Padua in Italy and another of the project leads.

But, he added, “they do suggest that they are not as robust as might have been assumed. More research is needed to identify whether there are conditions in which the unreplicated findings can be observed. Many Labs 2 suggests that diversity in samples and settings may not be one of them.”

Many Labs 2 represents scientists’ efforts to review reproducibility of research and find ways to make scientific findings more robust.

A pre-press version of the article, “Many Labs 2: Investigating Variation in Replicability Across Sample and Setting,” along with commentaries, appears [online](#).

###

For access to articles published in *Advances in Methods and Practices in Psychological Science*, please contact **Scott Sleek** at 202-293-9300 or ssleek@psychologicalscience.org.