

Is the Glass Half Empty or Half Full? Latest Results in the Replication Crisis in Psychology

TOMASZ WITKOWSKI

For years a chorus of voices has spoken of a replication crisis in psychology, particularly social psychology. In November 2018, the results of another large replication project called Many Labs 2 were published, which demonstrated that in half of known psychological experiments it is possible to generate results that are similar to—but slightly lower than—the original results. Over a period of several years, 186 researchers from sixty different laboratories on six different continents representing thirty-six different nationalities conducted replications of twenty-eight well-known studies. Over 7,000 participants from diverse cultures took part in each of them. By comparison, the average number of participants in the original studies was 112. The results demonstrated clearly that replicability of experiments is not dependent on cultural context. Brian Nosek from the Center for Open Science, who helped to coordinate the study, claims that their results could prove even more significant than those of the 2015 Reproducibility Project, in which a mere 36 percent of experiments were successfully reproduced.

Studies were selected for replication based on a wide range of criteria, including frequency of citations, demonstrating the impact of a given study on the whole of psychology. Another, more technical criterion was the possibility to conduct the study online via the internet, using a platform accessible to researchers from all sixty laboratories.

Among the experiments whose results could not be confirmed was one demonstrating the Lady Macbeth effect: like the character from the Shakespearean drama, haunted by memories of her own crimes and convinced that she could see traces of blood on her hands, participants asked to recall some behavior associated with a violation of

moral norms would have a strong tendency to engage in behaviors involving restoration and maintenance of physical cleanliness. Confirmation was also not achieved of results showing that people growing up with siblings are more altruistic than only children. There was no success in confirming the results of experiments showing that participants

Over 7,000 participants from diverse cultures took part in each replication study. The results demonstrated clearly that replicability of experiments is not dependent on cultural context.

performing tasks involving the use of words describing heat are more likely to believe in global warming than those who performed tasks linked with the use of words describing cold. Replication was also not achieved of the famous 2009 experiment by Yoel Inbar showing that people who had first experienced a feeling of aversion were later more homophobic than those from the control group.

However, there was successful replication of the results of a study by the Nobel Prize winner Daniel Kahneman, who together with Amos Tversky discovered the “framing effect,” in which people who are in possession of the same information make different decisions depending on the context in

which that data is presented to them. Many Labs 2 replicated the results of a study from 2007 on the so-called trolley dilemma. As in the original experiment, here the majority of participants considered it impermissible to shove one person onto the tracks in order to halt a wagon that could potentially kill five people. The trials of replication were also successfully handled by experiments on the primary attribution effect, the tendency to explain behaviors of observed persons in terms of internal and fixed categories (e.g., personality traits), while simultaneously discounting external situational factors.

As always when we achieve a 50/50 result, the trite question comes to mind of whether the glass is half full or half empty. The answer depends not only on our reserves of optimism or pessimism but also on the perspective we have adopted. If we remove our rose-tinted glasses and look at psychology, particularly social psychology, as a science that has for years struggled with methodological problems, beset by scandals surrounding numerous scientific fraudsters, our answer would, perhaps paradoxically, be positive. From this perspective 50 percent is quite a lot, as it means that a significant portion of the general rules discovered by psychology do not depend on cultural context (which is not to say that cultural context has no meaning in psychological experiments). Psychologists until quite recently have not had this certainty, for the majority of study results came from the WEIRDest (Western, educated, industrialized, rich, democratic) populations and mainly from American psychology students.

Another piece of good information coming from the results of Many Labs 2 is that a significant portion of our knowledge on human behavior is independent of the passage of time—the variable that is the bugaboo of so many

social scientists, whose work frequently turns out to be a snapshot of something ethereal. Perhaps this is the case of the study of altruism of only children and those raised with siblings, whose results were not confirmed by Many Labs 2. During the time from which the original study was conducted to the moment the replication was done, there were most certainly changes in the manner in which children are raised and changes in many cultures. Perhaps some functions that the home and family previously performed have been assumed by various institutions. This, in turn, could have led to contacts between only children and their peers being more extensive than in previous times, thus substituting for siblings. However, to determine whether these suspicions are accurate, it would be necessary to conduct a well-designed longitudinal study. The mere replication does not tell us.

The results of Many Labs 2 also reinforce convictions about the importance for study replicability of the effect size uncovered in the original study. Indeed, as it turned out, other factors such as the place of the study, order of tasks, or culture did not exert the same impact on replication results as did effect size. Either something could be repeated everywhere or not at all. We can draw clear conclusions for future research and publication work. The discoveries of the Many Labs 2 team, supplying invaluable pointers to contemporary researchers, give hope that the knowledge referred to by the nineteenth-century mathematician William Clifford as “safe truth” will expand in psychology, while we will learn about the replication crisis during classes on the history of psychology.

Before this happens, however, we should keep in mind that in the field of psychology almost 150,000 reviewed studies are published annually in English alone (144,291 in 2017 according to PsycINFO), making twenty-eight replicated studies over a period of several years a mere drop in the bucket. The perspective of scientists is also not the only one we should take into account when answering the question of how full our glass is. Entirely different

is the perspective adopted by the average user of psychological knowledge—practitioners applying knowledge from social psychology in management or in the shaping of social attitudes through advertisements or education. They will ask questions similar to those posed by users of other goods accessible to us. How many forged banknotes in circulation make a currency worthless? What percentage of false entries in Wikipedia would lead to its total uselessness as a source of information? What proportion of defective products will lead people to stop buying goods from a particular company? How much false information can be published by a newspaper before it loses the trust of its readers? Unfortunately, the results of Many Labs 2 mean that half of the pages in textbooks and other volumes used by researchers may contain false information. The problem is, we don't know which half.

Perhaps the inauthenticity of half of the banknotes in circulation is a problem for a currency, just as false information is a problem for the reliability of sources of information. That said, it does not appear that this applies to psychology. In 2018, the American Psychological Association reported that psychology presently enjoys greater popularity than in any other time in history. Every year around 1.6 million first-year students in America take an introduction to psychology course, and this number is growing. Forecasts for employment among psychologists predict a 19 percent increase in the need for psychologists by 2024. There is a similar level of interest in the study of psychology and services offered by psychologists in Europe. Is this growth in interest happening despite the replication crisis or perhaps because of it? Or perhaps both suppositions are true.

Tomasz Witkowski is a psychologist and writer. He is author of *Psychology Led Astray* and coauthor of *Psychology Gone Wrong*. He is the founder of the Polish Skeptics Club and specializes in debunking pseudoscience, particularly in the field of psychology and psychotherapy.