

# Association between contextual dependence and replicability in psychology may be spurious

Yoel Inbar<sup>a,1</sup>

The Reproducibility Project: Psychology (RP:P) attempted to replicate 100 cognitive and social-personality psychology studies (1). Van Bavel et al. (2) report an association ( $r = -0.23$ ,  $P = 0.024$ ) between a study's rated "contextual dependence" and whether the study was successfully replicated by the RP:P. However, this association is entirely the result of an omitted third variable: whether the study was in cognitive or social-personality psychology. Within each subdiscipline there is no relationship between context dependence and replicability:  $r = -0.08$ ,  $P = 0.54$  for social-personality psychology;  $r = -0.04$ ,  $P = 0.79$  for cognitive psychology (point-biserial correlations).

## The Third-Variable Problem

The RP:P coded studies as cognitive ( $n = 43$ ) or social-personality ( $n = 57$ ) psychology; the former were much more likely to replicate successfully [53% vs. 28%, by the standard Van Bavel et al. (2) used]. Cognitive psychology studies were also rated by Van Bavel et al. as less context-dependent than social-personality psychology studies ( $d = -1.85$ ,  $P < 0.001$ ). Therefore, the "third variable" of subdiscipline could explain the relationship between context-dependence and replicability. Van Bavel et al. are aware of this possibility, and report an analysis that they suggest addresses it: the interaction between subdiscipline and context dependence does not predict replicability ( $P = 0.877$ ). However, this analysis does not answer the question of interest. It asks "does the effect differ by subdiscipline?" not "is there an effect controlling for subdiscipline?"

To see the problem, suppose a researcher discovered a positive relationship between beardedness and height, but was concerned about a gender

confound: (i) men are more likely than women to have beards, and (ii) men are (on average) taller than women. Would it help the researcher to show that when considering men and women separately, there is no relationship in either group between beardedness and height (and thus no interaction)? It would not, but this is exactly what Van Bavel et al. (2) do. Van Bavel et al. report an effect of contextual sensitivity on replicability of odds ratio (OR) = 0.823 ( $P = 0.54$ ) for social-personality studies, and of OR = 0.892 ( $P = 0.78$ ) for cognitive studies (i.e., a null effect within each subdiscipline). Controlling for subdiscipline in a logistic regression (rather than testing the interaction), the previously significant effect of context dependence disappears, OR = 0.85,  $Z = -0.66$ ,  $P = 0.51$ . These null effects are not a result of restriction of range, as there is substantial variability in context sensitivity within each subdiscipline (Fig. 1).

## Conclusion

Van Bavel et al. (2) describe their data as showing that "contextual sensitivity appears to play an important role in replication success across multiple areas of psychology." A more apt summary is that contextual sensitivity is no longer associated with replicability once subdiscipline is taken into account.

Of course, cognitive psychology studies may be more replicable because they are less context-sensitive. However, cognitive and social-personality psychology differ in many other ways, such as greater use of within-subject designs in cognitive psychology. With Van Bavel et al.'s (2) data, it is simply not possible to say which—if any—of these differences are responsible for the observed differences in replicability between the two areas.

<sup>a</sup>Department of Psychology, University of Toronto, Toronto, ON, Canada M1C 1A4

Author contributions: Y.I. wrote the paper.

The author declares no conflict of interest.

Data deposition: A commented Stata do-file that reproduces all the analyses in this paper is posted at <https://dx.doi.org/10.6084/m9.figshare.3511120.v1>.

<sup>1</sup>Email: yoel.inbar@utoronto.ca.

