No evidence that experiment aversion is not a robust empirical phenomenon

Burcak Bas[a,1] (iD), Joachim Vosgerau[b] (iD), and Rachele Ciulli[c]

Mazar et al. (1) conducted a direct replication and six conceptual replications of the studies demonstrating experiment aversion (2, 3), the tendency to rate an experiment as less favorable than its treatment arms. Based on their findings, Mazar et al. conclude that experiment aversion "is neither generalizable nor robust" (1). We are questioning whether Mazar et al.'s findings justify this conclusion.

First, in the direct replication, the authors perfectly replicate experiment aversion, testifying to the replicability of experiment aversion. In that sense, experiment aversion constitutes a robust empirical finding.

In the six conceptual replications testing the generalizability of experiment aversion, the authors changed the original scenarios in a number of ways. For example, the authors employed six rather than one dependent variable to measure evaluations of hospitals that conduct experiments and instantiate treatment arms, asked participants to evaluate several hospitals rather than the decision of one doctor, described the hospitals' actions in the past rather than the present tense, etc. Unfortunately, these deviations from the original studies (2, 3) were not systematically and orthogonally manipulated, so it is not clear which changes contributed to not finding evidence of experiment aversion. Even more troubling, evidence for experiment aversion was found on some of the authors' measures but not on others.

While we agree that these findings suggest that experiment aversion is influenced by a host of factors, it is unclear what these factors are, what impact they have, and whether they are theoretically or practically interesting (4). One uninteresting possibility could be that the lengthy descriptions of the different hospitals that participants were asked to evaluate on six dimensions each reduced participants' attentiveness to the task, resulting in uninformative responses (5). Without systematically investigating the influence of each deviation from the original studies, it seems like little can be learned from the null findings. Concluding that experiment aversion is neither generalizable nor robust (1) is going far beyond Mazar et al.'s findings.

Author affiliations: ªDepartment of Marketing, WU Vienna (Vienna University of Economics and Business), Vienna 1020, Austria; ᵇDepartment of Marketing, Bocconi University, Milan 20136, Italy; and ᶜMarketing Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104

¹To whom correspondence may be addressed. Email: burcak.bas@wu.ac.at.

1. N. Mazar, C. T. Elbaek, P. Mitkidis, Experiment aversion does not appear to generalize. *Proc. Natl. Acad. Sci. U.S.A.* **120**, 1–5 (2023).
2. M. N. Meyer *et al.*, Objecting to experiments that compare two unobjectionable policies or treatments. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10723–10728 (2019).
3. P. R. Heck, C. F. Chabris, D. J. Watts, M. N. Meyer, Objecting to experiments even while approving of the policies or treatments they compare. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 18948–18950 (2020).
4. A. Zwaan, A. Etz, R. E. Lucas, M. B. Donnellan, Making replication mainstream. *Behav. Brain Sci.* **41**, e120 (2017).
5. T. Lenzner, L. Kaczmirek, A. Lenzner, Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Appl. Cogn. Psychol.* **24**, 1003–1020 (2010).