REPLY TO VELU AND IYER:

The promise and limits of "near-miss" pandemic-related research

Tim Johnson[a,1] and Dalton Conley[b,1]

Velu and Iyer (1) underscore the need and opportunity to study "near misses" in the COVID-19 era of social science research, thus elaborating on a key aspect of the causal inference challenges we discuss in our Opinion article (2). The literature on causal inference would classify "near misses" and "hypothetical histories" as "counterfactuals": depictions of what would have happened to an entity, absent some intervention, all else equal (3–5). Counterfactuals provide a baseline for measuring an intervention's causal effect. Theoretically, the difference between the observed outcome following an intervention and the result occurring had that intervention never happened constitutes the causal effect (4). However, since a researcher can only observe one state of the world (i.e., only the world in which the event occurred or the world in which it did not), a means of estimating the counterfactual must be found.

This challenge has vexed and will vex the COVID-19 era of social science research. Whereas the Vietnam Selective Service Lotteries that we use as an example in ref. 2 presented a clear counterfactual because of the random assignment of lottery numbers, the COVID-19 pandemic likely "intervened" nonrandomly such that folks (or communities) who suffered from COVID-19 probably differed from those who didn't in any number of ways—for instance, in their occupations and health—thus making the groups invalid counterfactuals. Any differences between outcomes may reflect the underlying differences of the groups, not the effect of COVID-19 itself.

In the face of this research problem, Velu and Iyer (1) remind us of an insightful way of thinking about how to search for and identify counterfactuals—look for near misses. Near misses involve instances in which some precipitating factor almost reached a value that triggered an event but failed to do so, and thus the event never intervened (e.g., the distance between planes diminishing but never reaching zero, and hence an accident never taking place). Comparing the outcomes of observations exposed to a near miss with those exposed to an event likely results in a comparison of similar entities, because only a very slight difference in a single variable's value distinguishes the entities. Indeed, this logic undergirds the regression discontinuity design—a widely used method of causal inference (6). For COVID-19 research, one could imagine comparing locations close to each other where arbitrary differences in transport led to just enough arriving travelers to produce an outbreak in one community while the other community had just slightly too few arriving travelers to create that problem.

Such research designs will result in valid causal inferences by relying on the type of thinking that Velu and Iyer (1) rightfully encourage. One caveat, however, concerns another challenge of COVID-19 social science research—generalizability (2). Comparing cases that differ subtly on the measure determining an event (or a near miss) entails looking at atypical cases and computing a local treatment effect (7), with limited generalizability. Still, it is preferable to make a limited, accurate claim rather than a broader false one; thus, analyzing "near misses" is an approach to pandemic-related research worth pursuing.

1 C. Velu, S. Iyer, Learning from near misses from COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2108269118 (2021).
2 D. Conley, T. Johnson, Opinion: Past is future for the era of COVID-19 research in the social sciences. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104155118 (2021).
3 S. L. Morgan, C. Winship, *Counterfactuals and Causal Inference* (Cambridge University Press, Cambridge, United Kingdom, ed. 2, 2015).
4 D. B. Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974).

[a]Atkinson Graduate School of Management, Willamette University, Salem, OR 97301; and [b]Department of Sociology, Princeton University, Princeton, NJ 08544

Author contributions: T.J. and D.C. wrote the paper.
The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: daltonclarkconley@gmail.com or tjohnson@willamette.edu.

Published September 27, 2021.

5 P. W. Holland, Statistics and causal inference. *J. Am. Stat. Assoc.* **81**, 945–960 (1986).
6 D. L. Thistlethwaite, D. T. Campbell, Regression-discontinuity analysis: An alternative to the ex post facto experiment. *J. Educ. Psychol.* **51**, 309–317 (1960).
7 G. W. Imbens, J. D. Angrist, Identification and estimation of local average treatment effects. *Econometrica* **62**, 467–475 (1994).