REPLY TO HAMAKER AND RYAN:

Within-sample temporal instability in cross-sectional estimates

Aaron J. Fisher[a,1], Bertus F. Jeronimus[b], and John D. Medaglia[c,d]

We offer our sincere gratitude to Hamaker and Ryan (1) for their comment. As we note in our manuscript (2), we believe the generalization issue in question to be a key problem in human subject research methodology. This challenge requires effort from all quarters to resolve. We also agree "that we should interpret measures in terms of what they are meant to represent." This is perhaps the most important point to be taken from our paper—that understanding individuals requires measuring and analyzing individuals. Thus, it may be fair to argue that comparisons between idiographic and nomothetic data structures offer limited value. To this end, we do not diverge from Hamaker and Ryan's position.

However, we should clarify that the repeated sampling paradigm we used explicitly avoids the problem of an asymptotically error-free estimate under infinite sampling conditions. As Hamaker and Ryan (1) note, our cross-sectional samples were not independent. Specifically, we resampled the same population, in its entirety, multiple times. As a result, the variability around the cross-sectional correlation estimate is the degree to which that estimate varies within the sampled population over time. Taken to the extreme, measuring a bivariate correlation in the entire human population at one time point would return a cross-sectional estimate without error. However, measuring the same bivariate correlation in the entire human population repeatedly would produce a distribution of estimates with a central tendency and a nonzero SD—the degree to which each measurement varied from the average across time points. Consequently, our estimate is not a proxy for the SE of the cross-sectional correlation estimate or for within-individual correlations, but a third representation, which estimates within-sample temporal instability in cross-sectional estimates.

As Hamaker and Ryan (1) would likely agree, this third representation is fundamentally not a representation of within-subject processes over time. Moreover, it is also fair to acknowledge that it likely does not reflect the true measure of the variability (i.e., temporal instability) of the cross-sectional estimate over time. Although the repeated sampling of any population will produce variation around the mean, some of this variation will no doubt be due to within-individual variance over time. This intrasubject variation might further include important stochastic, nonlinear, and phase information.

We believe that for group-derived statistical estimates to generalize to individuals, these estimates must hold true over time (and likewise correspond with temporal dynamics at the individual level). Given that human physiology and phenomenology unfold in time, quantifications of human experience must apply over repeated intervals. However, the variability that a single individual exhibits in behavior over repeated assessments is fundamentally different from the variability that a cross-sectional estimate exhibits from repeated testing. Thus, it may be inappropriate to use one estimate to represent or draw inferences about the other (3). Like Hamaker and Ryan (1), we advocate for statistical practices that distinguish group and individual effects, as these illustrate the difference between understanding people and claims about groups, which may or may not extrapolate to their constituent individuals.

[a]Department of Psychology, University of California, Berkeley, CA 94720; [b]Department of Developmental Psychology, Faculty of Behavioural and Social Sciences, Interdisciplinary Center Psychopathology and Emotion Regulation, Groningen University, 9712 TS Groningen, The Netherlands; [c]Department of Psychology, Drexel University, Philadelphia, PA 19104; and [d]Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104

Author contributions: A.J.F., B.F.J., and J.D.M. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: afisher@berkeley.edu.

1 Hamaker EL, Ryan O (2019) A squared standard error is not a measure of individual differences. *Proc Natl Acad Sci USA*, 10.1073/pnas.1818033116.
2 Fisher AJ, Medaglia JD, Jeronimus BF (2018) Lack of group-to-individual generalizability is a threat to human subjects research. *Proc Natl Acad Sci USA* 115:E6106–E6115.
3 Adolf JK, Fried E (2019) Ergodicity is sufficient but not necessary for group-to-individual generalizability. *Proc Natl Acad Sci USA*, 10.1073/pnas.1818675116.