



Reply to Krueger: Good point, wrong paper

Citation

Stanley, Damian A., Peter Sokol-Hessner, Mahzarin R. Banaji, and Elizabeth A. Phelps. 2011. Reply to Krueger: Good point, wrong paper. *Proceedings of the National Academy of Sciences* 108, no. 32: E411. doi:10.1073/pnas.1107937108.

Published version

<https://doi.org/10.1073/pnas.1107937108>

Link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:27416627>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

Reply to Krueger: Good point, wrong paper

In his letter, Krueger (1) raises an important issue regarding questions of prediction of behavior from tests of underlying mental states. We think it is important to make the distinction here between using the implicit association test (IAT) for diagnostic purposes (e.g., to classify individuals as 'racist') and using the IAT for research purposes (i.e., to investigate a potential contributing factor to trust behavior in our study) (2). We emphatically agree with Krueger (1) that, in its current instantiation, the IAT should not be used for diagnostic purposes. Indeed, the concern of Krueger (1) is exactly the one that we would raise ourselves if anyone attempted to use test scores to isolate individuals and hold them responsible. We have always explicitly eschewed such a notion. No test perfectly predicts behavior, including those tests that are often used for selection. The IAT, as currently implemented, is a research and educational technique, not a selection technique; although relevant when considering future uses of the IAT, the concerns of Krueger (1) do not apply here.

Our stated goal was to determine whether there was evidence of a behavioral link between two distinct psychological variables (disparity in trust behavior and implicit racial bias as measured by the IAT). We were motivated by evidence for a common neural substrate, but the fact that we found a correlation was not a foregone conclusion. We did not make the claim that one's IAT score was diagnostic of whether one was 'racist' in their trust behavior (2). Trust estimations are complex, and the evidence from our own study (2) and others suggested that they covaried with many factors, both explicit and implicit. We showed that the extent to which trust estimations varied with respect to a partner's race can be predicted in part by their IAT scores—in addition to and independently of other factors (2). As such, the correlation between one's IAT score and trust

behavior was an important indicator of the degree to which implicit attitudes may contribute to trust decisions.

The finding that, on average, there was little discriminatory trust behavior is not novel (3) and was supported by our original neuroimaging finding (4). In that study, there was little overall difference in the amygdala response to black vs. white faces but a significant correlation between implicit racial bias and amygdala activation. The presence of this correlation in the absence of a correlation between an explicit measure of race bias and amygdala activity was a theoretically important discovery.

We found a significant correlation between IAT test scores and trust behavior (2). This result meant that those individuals with stronger pro-white IAT scores were *more likely* to discriminate against blacks, *not* that they would definitely do so. Although Krueger (1) is right to worry about the potential misuse of the IAT in a diagnostic capacity, he is off-base in making that point about a paper (2) that showed, consistent with mounting evidence from dozens of studies (5), that implicit measures of attitudes and beliefs predicted ecologically realistic behavior.

Damian A. Stanley^a, Peter Sokol-Hessner^a, Mahzarin R. Banaji^b, and Elizabeth A. Phelps^{a,1}

^a*Department of Psychology, New York University, New York, NY 10003; and* ^b*Department of Psychology, Harvard University, Cambridge, MA 02138*

1. Krueger JI (2011) Decision-theoretic perspective on racial mistrust. *Proc Natl Acad Sci USA* 108:E410.
2. Stanley DA, Sokol-Hessner P, Banaji MR, Phelps EA (2011) Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proc Natl Acad Sci USA* 108:7710–7715.
3. Green AR, et al. (2007) Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *J Gen Intern Med* 22:1231–1238.
4. Phelps EA, et al. (2000) Performance on indirect measures of race evaluation predicts amygdala activation. *J Cogn Neurosci* 12:729–738.
5. Jost JT, et al. (2009) The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. *Res Organ Behav* 29:39–69.

Author contributions: D.A.S., P.S.-H., M.R.B., and E.A.P. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: liz.phelps@nyu.edu.