PSYCHOLOGICAL INQUIRY
2017, VOL. 28, NO. 4, 281–287
https://doi.org/10.1080/1047840X.2017.1373555

Routledge
Taylor & Francis Group

Reports of the Death of the Individual Difference Approach to Implicit Social Cognition May Be Greatly Exaggerated: A Commentary on Payne, Vuletich, and Lundberg

Benedek Kurdi and Mahzarin R. Banaji

Department of Psychology, Harvard University, Cambridge, Massachusetts

The sequential priming and lexical decision paradigms (Meyer & Schvaneveldt, 1971; Neely, 1976) first used to investigate the implicit social mind during the 1980s were firmly rooted in the study of learning and memory in experimental cognitive psychology and, as such, focused on the interpretation of group means to understand fundamental principles of mental processes. Researchers using the same methods to explore social cognition likewise employed these measures in within-participant designs to provide evidence for differential *average associations* between social categories (e.g., White vs. Black) and attributes (e.g., good vs. bad). For instance, Gaertner and McLaughlin (1983) demonstrated that participants were, *on average*, faster in making lexical decisions involving positive words following White primes. Fazio, Sanbonmatsu, Powell, and Kardes (1986) observed, *on average*, shorter response latencies to positive words following White primes and to negative words following Black primes. Dovidio, Evans, and Tyler (1986) showed that priming participants with the categories White or Black, *on average*, facilitated responding to traits stereotypically associated with the respective categories (also see Devine, 1989). The effects produced in experiments of this kind were small but highly robust and replicable across a range of different target groups and procedural variations (e.g., Banaji & Hardin, 1996; Blair & Banaji, 1996; Fazio, Jackson, Dunton, & Williams, 1995; Glaser & Banaji, 1999; Perdue & Gurtman, 1990; Wittenbrink, Judd, & Park, 1997, 2001).

The Individual Difference versus Experimental Approaches to Implicit Social Cognition

Because sequential priming paradigms produced small effects, they did not enable reliable measurement of individual differences. In fact, in 1995, Greenwald and Banaji saw "adapt[ing] these [i.e., implicit] methods for efficient assessment of individual differences in implicit social cognition" as "perhaps the most significant remaining challenge" (p. 20) for the field. Three years later, a step was made toward addressing this challenge with the appearance of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), which yielded sufficiently large effect sizes to allow for meaningful estimates of individual differences in implicit attitudes and stereotypes. Since then, countless studies have used the IAT and its variants, including the Go/No-Go Association Task (Nosek & Banaji, 2001), the personalized IAT (Olson & Fazio, 2004), the Implicit Relational Assessment Procedure (Barnes-Holmes et al., 2006), the single-category IAT (Karpinski & Steinman, 2006), and the brief IAT (Sriram & Greenwald, 2009), as well as the Affect Misattribution Procedure (Payne, Cheng, Govorun, & Stewart, 2005) to predict individual differences in behavior (for meta-analyses, see Cameron, Brown-Iannuzzi, & Payne, 2012; Greenwald, Poehlman, Uhlmann, & Banaji, 2009a; Kurdi et al., 2017; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013).

Studies using the *individual difference approach* to social cognition—that is, probing the relationship between measures of implicit attitudes and stereotypes, on one hand, and measures of behavior, on the other hand—have been numerous and highly heterogeneous. Investigators have used implicit race attitudes to predict widely divergent outcome measures, including seating distance (Amodio & Devine, 2006), the categorization of racially ambiguous faces (Hugenberg & Bodenhausen, 2004), voting intentions and voting behavior (Caruso, Mead, & Balcetis, 2009; Greenwald, Smith, Sriram, Bar-Anan, & Nosek, 2009b; Plant et al., 2009), attributions of guilt (Levinson, Cai, & Young, 2010; Mastro, Lapinski, Kopacz, & Behm-Morawitz, 2009; Perugini, O'Gorman, & Prestwich, 2007), and even tobacco use (Krieger et al., 2011).

Parallel to the individual difference approach, the past two decades have also seen the emergence of an entirely different and almost completely disjoint literature using the *experimental approach* to implicit social cognition, which has documented myriad ways in which contextual influences can modulate implicit social attitudes and stereotypes (for reviews, see Blair, 2002; Lai, Hoffman, & Nosek, 2013). Some early examples for this line of research include studies demonstrating immediate changes in implicit attitudes as a result of different experimental manipulations such as self-image enhancement (Sinclair & Kunda, 1999), social influence

CONTACT Benedek Kurdi kurdi@g.harvard.edu Department of Psychology, Harvard University, William James Hall 1580, 33 Kirkland Street, Cambridge, MA 02138, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hpli.

The R code used to conduct the simulations and to prepare the figures included in this article is available for download from the Open Science Framework (https://osf.io/y6pzm/).

(Lowery, Hardin, & Sinclair, 2001), implementation intentions (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000), exposure to counterstereotypic exemplars (Dasgupta & Greenwald, 2001), and recategorization (Mitchell, Nosek, & Banaji, 2003). Lai et al. (2014) conducted a large-scale comparative investigation of different interventions for shifting implicit racial attitudes toward neutrality and found significant effects for vivid imagery, shifting group boundaries through competition or threat, practicing the IAT with counterstereotypic exemplars, and different implementations of evaluative conditioning.

Thus, research using the individual difference approach has viewed implicit social attitudes as relatively stable traits of individuals, whereas research using the experimental approach has focused on the effects of the immediate context on implicit social attitudes. Given their completely different fundamental assumptions, these two areas of research have unfolded along parallel tracks, with little cross-fertilization or even cross-communication. Experimental studies have treated individual variation as a source of error that should be reduced using the strongest possible manipulations; correlational studies have largely ignored the evidence on contextual variation.

Putting One and One Together

The theoretical piece by Payne, Vuletich, and Lundberg (this issue) takes a crucial step toward resolving the problem of separation between experimental and correlational approaches to implicit social cognition. Payne and colleagues forcefully point out the curious state of affairs that has persisted over the past two decades in which one group of studies has treated implicit attitudes as reflecting individual variation with some negligible error variance accounted for by contextual differences, whereas a different group of studies has treated implicit attitudes as reflecting contextual variation with some negligible error variance accounted for by individual differences. At least without some qualifications, both sets of assumptions cannot be accurate at the same time.

The solution offered by Payne and colleagues is the following. They observe, accurately, that the evidence on the malleability of implicit attitudes and stereotypes in the face of contextual influences is overwhelming (Blair, 2002; Lai et al., 2014). They observe, also accurately, that the evidence on the relationship between implicit measures of social attitudes and stereotypes, on one hand, and measures of intergroup behavior, on the other, has been mixed, with meta-analytic investigations producing small, or at most small-to-medium, effect sizes. Estimates of the relationship between implicit and criterion measures observed at the level of individual participants (*individual implicit–criterion correlation*; iICC) include $r = .24$ ($k = 32$) in the race domain, $r = .18$ ($k = 15$) for gender and sexual orientation, and $r = .20$ ($k = 15$) in other intergroup contexts (Greenwald et al., 2009a), $r = .15$ ($k = 36$) for race and $r = .12$ ($k = 14$) for ethnicity (Oswald et al., 2013), and $r = .28$ ($k = 30$) across all intergroup contexts (Cameron et al., 2012). To date, the evidence appears to be unambiguous: Variations in context have a demonstrated ability to bring about considerable change in implicit attitudes, whereas individual differences in implicit attitudes account for a measly 1%, or at most 8%, of the variance in intergroup discrimination. It seems that the time has come to issue a death certificate for the individual difference approach to implicit social cognition.

At the same time, Payne and colleagues highlight a growing body of work that suggests that even though individual differences in implicit intergroup cognition may not do much to explain individual differences in intergroup behavior, correlational studies may not be entirely doomed provided that investigators shift the unit of analysis from individuals to larger geographic units. Studies on *aggregate implicit–criterion correlations* (aICCs) have generated some remarkable findings, including the following: (a) More pervasive implicit gender stereotypes aggregated at the national level show a strong association with larger gender gaps in mathematics and science achievement, also measured at the national level (Nosek et al., 2009); (b) more negative implicit attitudes toward overweight individuals, as measured at the national level, show a strong association with higher average levels of obesity (Marini et al., 2013); (c) more negative implicit race attitudes among White Americans, as measured at the county level, show a strong association with a higher percentage of African American residents (Rae, Newheiser, & Olson, 2015); (d) more positive implicit attitudes toward Arabs, as measured at the city level in France, show a strong association with residents' willingness to participate in public demonstrations against racism following the attack on *Charlie Hebdo* staff (Zerhouni, Rougier, & Muller, 2016); (e) more negative implicit attitudes toward White Americans among African Americans, as measured at the county level, show a strong association with higher death rates among African Americans (Leitner, Hehman, Ayduk, & Mendoza-Denton, 2016); and (f) more negative implicit race attitudes among White Americans, as measured at the level of core-based statistical areas, show a strong association with disproportionate use of lethal force against African Americans by police officers (Hehman, Flake, & Calanchini, 2017).

A crucial methodological aspect that unifies these successful aICC prediction studies is that implicit attitudes or stereotypes exhibited by participants from the Project Implicit educational website (http://implicit.harvard.edu) were aggregated within a geographic area and used to predict criterion measures that were obtained from the same area but not the same group of participants. Rather, criterion behaviors were performed by high school students (Nosek et al., 2009), police officers (Hehman et al., 2017), or the general population. According to Payne and colleagues, these findings can also be accounted for within the framework of the experimental approach. Contextual influences, such as the manipulations used in experimental studies, make certain associations or propositions (e.g., Black–good or "Black people are good") temporarily more accessible to participants. However, such influences are not limited to experimental contexts; rather, they are also pervasive under ecologically valid conditions. Moreover, the spatial distribution of contextual influences in everyday life is far from random: Individuals living in the same geographic area tend to be exposed to similar contextual influences, including physical spaces, content distributed by the media, commonly used symbols, social networks, and language (Caliskan, Bryson, & Narayanan, 2017; Shepherd, 2011). Therefore, over time, contextual influences accumulate and individuals living in physical proximity to one another become similar to one another in terms of the associations and propositions accessible to them.

Idiosyncratic individual variation may exist, but it will be negligible compared to these aggregate effects.[1]

It May Be Premature to Declare the Individual Difference Approach Dead

We are in full agreement with Payne and colleagues regarding the importance of integrating correlational and experimental approaches to implicit social cognition, as well as about their assessment of the literatures on (a) experimental change to implicit attitudes and (b) the relationship between implicit attitudes and behavior at the level of geographic areas rather than individuals. However, it should be noted that individual versus contextual approaches to implicit social cognition need not be a zero-sum game. The aICCs (i.e., associations between implicit and criterion measures at the level of geographic areas) do not reveal anything about the iICCs (i.e., the same associations at the level of individuals) and vice versa. In fact, whereas the aICC between body mass index and implicit weight bias at the national level is *positive*, the iICC between the same measures is *negative*, that is, heavier individuals tend to exhibit *lower* levels of implicit weight bias (Marini et al., 2013).

Some findings from our recent comprehensive meta-analysis of studies using the Implicit Association Test and its variants to predict intergroup discrimination (Kurdi et al., 2017) indicate that, in spite of the small overall meta-analytic effect sizes just mentioned, it may be premature to declare the individual difference approach to implicit social cognition dead. We think that the jury should sit out on this issue because the extant evidence on iICCs may not provide an accurate estimate of the population effect size. We believe that this is true for the following reasons, each of which is discussed in more detail next: (a) inappropriate treatment of measurement error, (b) low levels of statistical power, (c) high degree of variability in effect sizes accounted for by methodological differences across studies, and (d) a lack of consideration of the theoretical relationship between implicit attitudes and behaviors as well as between manifest behaviors and the underlying constructs that they are intended to measure.

Measurement Error

We agree with Payne and colleagues about the noteworthy nature of the high aICCs observed in a number of recent studies. However, it is not entirely clear that lower iICCs are the accurate reflection of a smaller population effect size. In this section we consider the possibility that the discrepancy between aICCs and iICCs may, at least in part, be due to the fact that *the two designs treat measurement error in fundamentally different ways.*

It is a well-known psychometric fact that using noisy manifest measures of underlying latent constructs (such as using an IAT D score to measure implicit attitudes) places an upper bound on the strength of the statistical relationship found between those manifest variables in any given sample (e.g., Furr & Bacharach, 2013). We use Monte Carlo simulations conducted in the R statistical computing environment (https://osf.io/y6pzm/) to show that as larger numbers of individuals are sampled from a geographic unit, the more the obtained aICC approaches the population parameter because the effects of measurement error become increasingly more inconsequential (see Figure 1).[2] In this setup, individuals are nested within geographic units and shared membership in geographic units explains some variance in implicit attitudes and criterion behaviors, with all other sources of variance, including measurement error and individual differences, treated as error variance.

From a statistical perspective, the situation is similar in the case of iICCs. Measurements of implicit attitudes and behavior are nested within individuals and individual differences explain some variance in observed implicit attitudes and criterion behaviors. *Crucially, if the effects of measurement error on the sample correlation are to become negligible, multiple measurements of implicit attitudes and behavior must be obtained from each individual.* Statistically, this parallels the observation made earlier that in order to arrive at an accurate estimate of aICCs, multiple individuals must be sampled from each geographic unit. As shown by a second Monte Carlo simulation, assuming a population correlation of $\rho = .4$ and reliabilities of $R_{xx} = .73$ for implicit measures and $R_{yy} = .79$ for criterion measures,[3] one obtains a sample correlation of $r = .30$ if only one indicator of implicit cognition and one indicator of behavior are used, $r = .35$ for a composite of two indicators each, and $r = .38$ for a composite of five indicators each. That is, by using multiple indicators of implicit attitudes and behavior, studies would be able to achieve considerably higher iICCs (see Figure 2). The estimates are similar if one assumes a more realistic criterion reliability of $R_{yy} = .50$.

However, critically, *most iICC studies to date have correlated a single measure of implicit attitude with a single measure of behavior*. Given that both attitude and criterion measures contain a nontrivial portion of error variance, the estimates of the relationship obtained on the basis of single measures remain considerably below the population value. This situation is equivalent to hypothetical aICC studies sampling a very small number of individuals from each geographic unit included. The large difference in predictive power between the individual and aggregate levels may disappear or at least shrink considerably if iICC studies used multiple indicators of implicit cognition and behavior and

[1] In addition to low iICCs along with high aICCs, Payne and colleagues raise two additional findings that seem puzzling from the perspective of the individual difference approach: (a) temporally stable mean levels of implicit attitudes and (b) developmental invariance in implicit attitudes, both in spite of low test–retest correlations. Addressing these issues in detail would be beyond the scope of the present commentary. Nevertheless, it should be pointed out that the interpretation of low test–retest correlations, especially over the course of multiple weeks, is ambiguous. It is conceivable that, as assumed by Payne et al., relatively low test–retest correlations are due to genuine malleability in implicit attitudes rather than to the properties of implicit measures. However, this is certainly not the only possibility. Test–retest correlations could be lower than split-half correlations because certain sources of measurement error may operate across but not within administrations of the test (Viswanathan, 2005).

[2] In the simulation, the assumed population correlation between the implicit and criterion measure is $\rho = .4$, and geographic units are assumed to account for 3% of the variance in both implicit and criterion measures. The exact assumed values are arbitrary and do not alter the substantive argument.

[3] $R_{xx} = .73$ and $R_{yy} = .79$ are the mean reliability values observed by Kurdi et al. (2017). However, there is reason to assume that this value overestimates the reliability of criterion measures due to selective reporting. That is, authors tended to report estimates of internal consistency for well-established and validated instruments used as criterion measures but not for ad hoc measures of behavior.

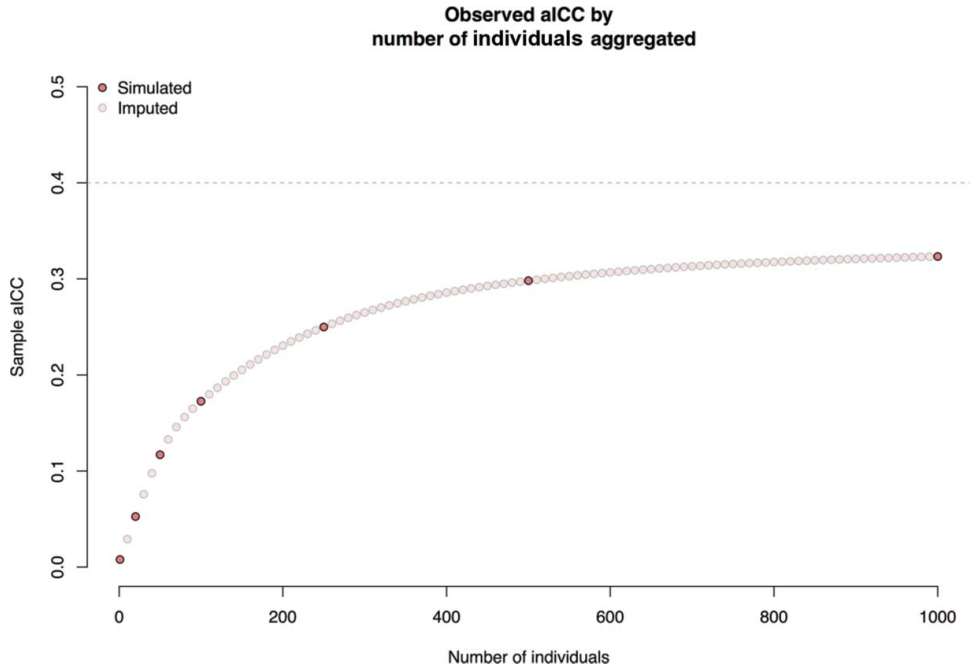

Figure 1. Results of a Monte Carlo simulation showing the effect of the number of individuals aggregated within each geographic unit on the observed aggregate implicit–criterion correlation (aICC). *Note*: Simulated effect sizes are shown in dark red and effect sizes imputed using a smoothing spline for the sake of completeness are shown in light red. The dashed gray line marks the population effect size ($\rho = .4$).

relied on analytic approaches that explicitly take into account measurement error (Westfall & Yarkoni, 2016).

Statistical Power

A second limitation of iICC prediction studies is woefully *low levels of statistical power*. In fact, in our recent meta-analytic investigation, we found a median sample size of 78, which corresponds to a mean level of 29% post hoc power, even among the subset of studies with a self-declared primary focus on predicting individual differences in intergroup behavior from individual differences in implicit cognition. Such low sample sizes are problematic because underpowered studies fail to provide meaningful evidence about the underlying

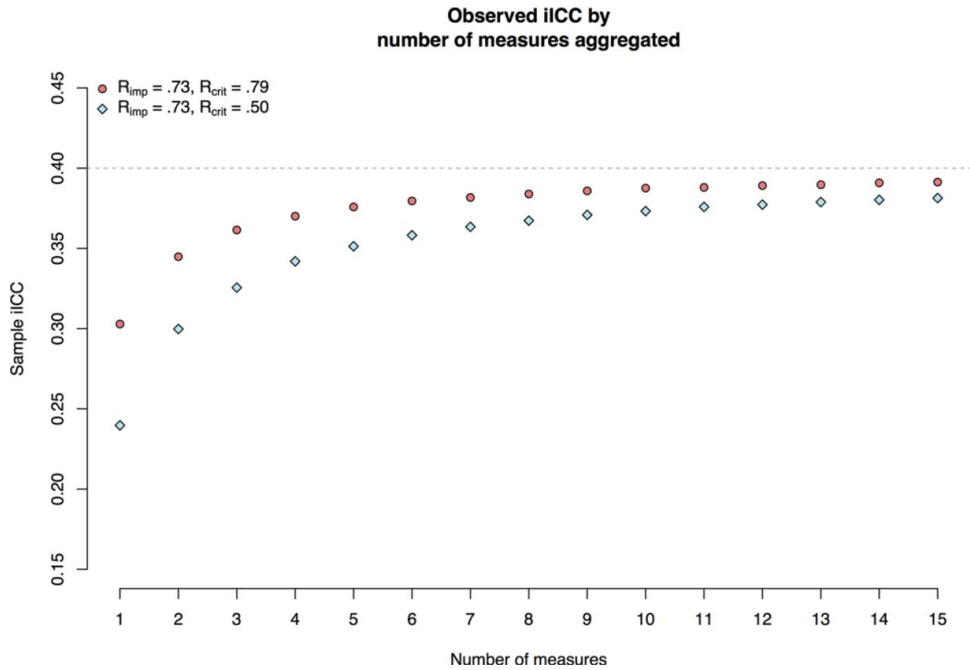

Figure 2. Results of two Monte Carlo simulations showing the effect of the number of measures aggregated within each individual on the obtained individual implicit–criterion correlation (iICC). *Note*: Red dots show simulated effect sizes assuming a reliability value of $R = .73$ for implicit measures and $R = .79$ for criterion measures, and blue diamonds show effect sizes assuming a reliability value of $R = .73$ for implicit measures and $R = .50$ for criterion measures. The dashed gray line marks the population effect size ($\rho = .4$).

population effect size and lead to low rates of generalizability and reproducibility (Button et al., 2013; Cohen, 1992; Higginson & Munafò, 2016).

Methodological Variation

A third factor that makes the meta-analytic estimate of iICCs of limited utility is the *high degree of heterogeneity among effect sizes*, some of which can be accounted for by a number of relatively simple methodological factors, including (a) the *type of implicit measure* used, with the standard IAT (Greenwald et al., 1998) producing larger effect sizes than its variants (Karpinski & Steinman, 2006; Nosek & Banaji, 2001; Olson & Fazio, 2004; Sriram & Greenwald, 2009); (b) *attribute polarity*, that is, the extent to which the attributes used on the implicit measure are polar opposites of each other, with high-polarity attributes such as *good–bad* or *strong–weak* producing larger effects than low-polarity attributes such as *smart–athletic* or *American–Muslim*; (c) the *type of criterion measure* used, with relative criterion measures (e.g., donation to a Black student group compared to a White student group) producing larger effects than absolute criterion measures (e.g., donation to a Black student group); and (d) *correspondence* between the implicit and criterion measure, that is, the extent to which implicit and criterion measures are compatible with each other (Ajzen & Fishbein, 1977), with high-correspondence implicit–criterion pairs producing larger effect sizes.

The variability in the magnitude of iICCs due to such methodological moderators, and thus the sample size required to ensure adequate statistical power, is substantial. Based on multivariate metaregressions, a study using a Go/No-Go Association Task with low-polarity attributes, an absolute measure of behavior, and low correspondence between the two is expected to produce an iICC of $r = .013$, whereas a study using a standard IAT with high-polarity attributes, a relative measure of behavior, and high correspondence between the two is expected to produce an iICC of $r = .242$. For the former, 48,999 participants would be necessary to achieve adequate (80%) power; for the latter, 130 participants are sufficient. Given extremely small samples, studies making methodological choices associated with smaller expected effect sizes cannot form the basis of valid inferences about the extent to which individual differences in implicit cognition predict individual differences in discriminatory behavior.

Validity and Reliability of Behavioral Measures

Implicit and explicit measures of cognition have been scrutinized for their psychometric properties and construct validity for decades (e.g., Banaji, 2001; Bar-Anan & Nosek, 2013; Cunningham, Preacher, & Banaji, 2001; Gawronski, Morrison, Phills, & Galdi, 2017; Thurstone, 1928). By contrast, as pointed out by multiple recent reviews of the relationship between attitudes and behavior, most criterion measures used in prediction studies tend to be ad hoc, with very little information about their construct validity or reliability (Carlsson & Agerström, 2016; Talaska, Fiske, & Chaiken, 2008). For instance, the studies included in our own-meta analysis tended to involve (a) single-shot behaviors, (b) undergraduate participants observed in a laboratory setting, and (c) conditions under which the purpose of the study was highly transparent to participants. These characteristics of the modal iICC study stand in sharp contrast to the aICC studies just reviewed, which tended to involve (a) repeated behaviors or behaviors aggregated across multiple individuals, (b) heterogeneous samples of participants observed in real-world settings, and (c) no transparency whatsoever in terms of the purpose of the study, especially with regard to criterion behaviors. By more carefully considering these and other theoretically relevant aspects of the behavioral measures used and, to the extent possible, more closely approximating aggregate-level studies, individual-level prediction studies may produce higher correlations between implicit and criterion measures.

Looking to the Future

By highlighting the necessity for integration across the currently disjoint experimental and individual difference traditions, the article by Payne et al. (this issue) constitutes a major advance in our understanding of implicit social cognition. Measures of implicit cognition cannot (a) largely reflect contextual influences with negligible effects of individual differences, as assumed by the experimental approach, and at the same time (b) largely reflect individual differences with negligible context effects, as assumed by the individual difference approach. Because of the contribution by Payne and colleagues, this glaring contradiction, which so far has remained largely unaddressed in the literature, has finally been given the attention that it deserves. In addition, by pointing out large correlations between measures of implicit social cognition and intergroup behavior as aggregated at the level of geographic units, Payne and colleagues provide a welcome reminder that the individual is not the only possible unit of analysis in considering the predictive validity of implicit measures. Beyond its theoretical implications, this novel perspective also suggests that the practical significance of implicit social cognition in influencing societally relevant outcomes need not hinge on correlations between implicit attitudes and behaviors measured at the individual level.

In this commentary we have argued that, pending major improvements in study design, it may be premature to conclude that implicit measures are poor predictors of behavior at the individual level and thus that implicit measures are primarily a reflection of situations rather than individuals. Individual-level prediction studies may be able to obtain correlations that are similar in magnitude to those found in aggregate-level studies if they (a) used multiple indicators of implicit attitudes and behavior, (b) used larger samples and thus had better statistical power, (c) took into account the effects of methodological choices on the magnitude of the correlations obtained, and (d) more carefully considered the theoretical relationship between implicit attitudes and behavior, as well as between manifest behaviors and the underlying constructs that they are intended to measure. Unless and until a large number of studies implement these proposed changes, conclusions about the relationship between implicit and behavioral measures at the individual level and, more generally, the merits of the individual

difference approach to the study of implicit social cognition will be of limited validity.

Current evidence is insufficient to discount the possibility that individual differences have some nontrivial influence on implicit social cognition. In fact, differences in implicit attitudes, as measured involving the same individuals across multiple occasions, may arise because of multiple components, including relatively stable components due to (a) individual differences and (b) chronic contextual effects as well as (c) a malleable component that reflects effects of the immediate situation. In the future, a better understanding of the processes by which implicit social cognition emerges from an interplay of persons and situations may be achieved by employing designs that systematically investigate the relative contributions of individual and contextual factors to implicit attitudes, for instance, within the framework of generalizability theory (Brennan, 2001). Moreover, the effects of person and situation may not necessarily be additive, making multilevel approaches potentially desirable: Two individuals who have the same, transsituationally stable levels of implicit attitude may behave quite differently given different immediate situational affordances and, conversely, two individuals with different transsituationally stable levels of implicit attitude may behave the same provided that the situation in which they are placed is sufficiently powerful.

Once the proper tests have been conducted, and they reveal that individual-level correlations between measures of implicit social cognition and intergroup behavior are, in fact, negligible compared to aggregate-level correlations, we will join Payne et al. in declaring the individual difference approach to implicit social cognition to be dead.

ORCID

Benedek Kurdi http://orcid.org/0000-0001-5000-0584
Mahzarin R. Banaji http://orcid.org/0000-0002-5941-7455

References

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918. doi:10.1037/0033-2909.84.5.888

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*(4), 652–661. doi:10.1037/0022-3514.91.4.652

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger III, J. S. Nairne, & I. Neath (Eds.), *The nature of remembering: Essays in Honor of Robert G. Crowder* (pp. 117–149). Washington, DC: American Psychological Association.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, *7*(3), 136–141. doi:10.1111/j.1467-9280.1996.tb00346.x

Bar-Anan, Y., & Nosek, B. A. (2013). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods*, *46*(3), 668–688. doi:10.3758/s13428-013-0410-6

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*(3), 242–261. doi:10.1207/S15327957PSPR0603_8

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*(6), 1142–1163. doi:10.1037/0022-3514.70.6.1142

Brennan, R. L. (2001). *Generalizability theory*. New York, NY: Springer. doi:10.1007/978-1-4757-3456-0

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. doi:10.1038/nrn3475

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. doi:10.1126/science.aal4230

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, *16*(4), 330–350. doi:10.1177/1088868312440047

Carlsson, R., & Agerström, J. (2016). A closer look at the discrimination outcomes in the IAT literature. *Scandinavian Journal of Psychology*, *57*(4), 278–287. doi:10.1111/sjop.12288

Caruso, E. M., Mead, N. L., & Balcetis, E. (2009). Political partisanship influences perception of biracial candidates' skin tone. *Proceedings of the National Academy of Sciences*, *106*(48), 20168–20173. doi:10.1073/pnas.0905362106

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. doi:10.1037/0033-2909.112.1.155

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *12*(2), 163–170. doi:10.1111/1467-9280.00328

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*(5), 800–814. doi:10.1037/0022-3514.81.5.800

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. doi:10.1037/0022-3514.56.1.5

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, *22*(1), 22–37. doi:10.1016/0022-1031(86)90039-9

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027. doi:10.1037/0022-3514.69.6.1013

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238. doi:10.1037/0022-3514.50.2.229

Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: An introduction*. Los Angeles, CA: Sage.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly*, *46*(1), 23–30. doi:10.2307/3033657

Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures. *Personality and Social Psychology Bulletin*, *43*(3), 300–312. doi:10.1177/0146167216684131

Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology*, *77*(4), 669–687. doi:10.1037/0022-3514.77.4.669

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27. doi:10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. doi:10.1037/0022-3514.74.6.1464

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009a). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. doi:10.1037/a0015575

Greenwald, A. G., Smith, C. T., Sriram, N., Bar-Anan, Y., & Nosek, B. A. (2009b). Implicit race attitudes predicted vote in the 2008 U.S.

presidential election. *Analyses of Social Issues and Public Policy*, *9*(1), 241–253. doi:10.1111/j.1530-2415.2009.01195.x

Hehman, E., Flake, J. K., & Calanchini, J. (2017). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science, 2*, 1948550617711229–9. doi:10.1177/1948550617711229

Higginson, A. D., & Munafò, M. R. (2016). Current incentives for scientists lead to underpowered studies with erroneous conclusions. *PLoS Biology*, *14*(11), e2000995–14. doi:10.1371/journal.pbio.2000995

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization. The role of prejudice and facial affect in race categorization. *Psychological Science*, *15*(5), 342–345. doi:10.1111/j.0956-7976.2004.00680.x

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*(1), 16–32. doi:10.1037/0022-3514.91.1.16

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*(5), 871–888. doi:10.1037/0022-3514.78.5.871

Krieger, N., Waterman, P. D., Kosheleva, A., Chen, J. T., Carney, D. R., & Smith, K. W., … Samuel, L. (2011). Exposing racial discrimination: Implicit & explicit measures—The *My Body, My Story* study of 1005 US-born Black & White community health center members. *PLoS ONE*, *6*(11), e27636. doi:10.1371/journal.pone.0027636

Kurdi, B., Seitchik, A. E., Axt, J. A., Carroll, T. J., Karapetyan, A., Kaushik, N. ... Banaji, M. R. (2017). *Predicting intergroup discrimination using the Implicit Association Test: Systematic review, meta-analysis, and recommendations for future research*. Manuscript submitted for publication.

Lai, C. K., Hoffman, K. M., & Nosek, B. A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, *7*(5), 315–330. doi:10.1111/spc3.12023

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., & Joy-Gaba, J. A., … Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*(4), 1765–1785. doi:10.1037/a0036260

Leitner, J. B., Hehman, E., Ayduk, O., & Mendoza-Denton, R. (2016). Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Social Science & Medicine*, *170*(C), 220–227. doi:10.1016/j.socscimed.2016.10.007

Levinson, J. D., Cai, H., & Young, D. M. (2010). Guilty by implicit racial bias: The guilty/not guilty Implicit Association Test. *Ohio State Journal of Criminal Law*, *8*(1), 187–208.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, *81*(5), 842–855. doi:10.1037/0022-3514.81.5.842

Marini, M., Sriram, N., Schnabel, K., Maliszewski, N., Devos, T., Ekehammar, B., … Nosek, B. A. (2013). Overweight people have low levels of implicit weight bias, but overweight nations have high levels of implicit weight bias. *PLoS ONE*, *8*(12), e83543. doi:10.1371/journal.pone.0083543

Mastro, D., Lapinski, M. K., Kopacz, M. A., & Behm-Morawitz, E. (2009). The influence of exposure to depictions of race and crime in TV news on viewers' social judgments. *Journal of Broadcasting & Electronic Media*, *53*(4), 615–635. doi:10.1080/08838150903310534

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. doi:10.1037/h0031564

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*(3), 455–469. doi:10.1037/0096-3445.132.3.455

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, *4*(5), 648–654. doi:10.3758/BF03213230

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–644. doi:10.1521/soco.19.6.625.20886

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597. doi:10.1073/pnas.0809921106

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, *86*(5), 653–667. doi:10.1037/0022-3514.86.5.653

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–192. doi:10.1037/a0032734

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293. doi:10.1037/0022-3514.89.3.277

Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology*, *26*(3), 199–216. doi:10.1016/0022-1031(90)90035-K

Perugini, M., O'Gorman, R., & Prestwich, A. (2007). An ontological test of the IAT: Self-activation can increase predictive validity. *Experimental Psychology*, *54*(2), 134–147. doi:10.1027/1618-3169.54.2.134

Plant, E. A., Devine, P. G., Cox, W. T. L., Columb, C., Miller, S. L., Goplen, J., & Peruche, B. M. (2009). The Obama effect: Decreasing implicit prejudice and stereotyping. *Journal of Experimental Social Psychology*, *45*(4), 961–964. doi:10.1016/j.jesp.2009.04.018

Rae, J. R., Newheiser, A.-K., & Olson, K. R. (2015). Exposure to racial out-groups and implicit race bias in the United States. *Social Psychological and Personality Science*, *6*(5), 535–543. doi:10.1177/1948550614567357

Shepherd, H. (2011). The cultural context of cognition: What the Implicit Association Test tells us about how culture works. *Sociological Forum*, *26*(1), 121–143. doi:10.1111/j.1573-7861.2010.01227.x

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*(5), 885–904. doi:10.1037/0022-3514.77.5.885

Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology*, *56*(4), 283–294. doi:10.1027/1618-3169.56.4.283

Talaska, C. A., Fiske, S. T., & Chaiken, S. (2008). Legitimating racial discrimination: Emotions, not beliefs, best predict discrimination in a meta-analysis. *Social Justice Research*, *21*(3), 263–296. doi:10.1007/s11211-008-0071-2

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, *33*(4), 529–527. doi:10.1086/214483

Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks, CA: Sage.

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS ONE*, *11*(3), e0152719. doi:10.1371/journal.pone.0152719

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*(2), 262–274. doi:10.1037/0022-3514.72.2.262

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology*, *37*(3), 244–252. doi:10.1006/jesp.2000.1456

Zerhouni, O., Rougier, M., & Muller, D. (2016). "Who (Really) is Charlie?" French cities with lower implicit prejudice toward arabs demonstrated larger participation rates in Charlie Hebdo Rallies. *International Review of Social Psychology*, *29*(1), 69–76. doi:10.5334/irsp.50