Running Head: RESEARCH ON THE ATTITUDE-BEHAVIOR RELATION

It's Time to Do More Research on the Attitude-Behavior Relation: A Commentary on Implicit Attitude Measures

Perspective

Wenhao Dai

University of Illinois at Urbana–Champaign

Dolores Albarracin

University of Pennsylvania

Correspondence may be addressed to Dolores Albarracin, dalba@upenn.edu, 202 S. 36th St, Philadelphia PA 19104. Wenhao Dai, wdai7@illinois.edu, 603 E. Daniel St., Champaign, IL 61822

Abstract

The recent exchange about implicit attitudes is an acute reminder of the need to pay research attention to the correlation between implicit attitudes and overt behavior. Current implicit measures are excellent to detect evaluatively relevant associations arising from specific and variable internal states and predict judgments when people lack the motivation and ability to control those judgments. However, there is no convincing evidence of a strong correlation between such implicit attitudes and overt behavior when people's ability and motivation to control the influence of these attitudes is low. Researchers should improve implicit measures by better integrating action, target, level, and context into the measurement procedures and then reexamine if these improved measures predict socially undesirable behaviors when ability and motivation to control behavior are low.

Keyword: Attitude behavior relation; implicit attitudes; attitude behavior correspondence

It's Time to Do More Research on the Attitude-Behavior Relation: A Commentary on the Implicit Attitude Research and Implicit Measures

Implicit attitudes and implicit measures of attitudes have received research attention both within and outside the field of psychology for several decades. Brownstein, Madva, and Gawronski's interesting exchange with Machery presents two distinct points of view about implicit attitude measures and their contributions to attitude research. Brownstein et al. (2019) view implicit measures as a useful psychological instrument and have optimistic expectations for their future, whereas Machery (2021) presents a pessimistic assessment of the reliability and validity of these measures. Specifically, Machery highlights what he describes as the "anomalies" of implicit measures, including low validity, low reliability, low predictive power, and limited casual efficacy. These "anomalies" are addressed point-by-point in a follow-up response from Gawronski, Brownstein, and Madva (2022), who conclude that each "anomaly" is to be expected because implicit measures capture transient states instead of traits. Machery (2022) responds unconvinced, highlighting that the field of implicit attitudes has failed to accomplish its initial goal of measuring hidden attitudes in a reliable and valid way.

This debate is not new, as similar points have been raised before (Schimmack, 2021; Arkes & Tetlock, 2004; Tetlock & Mitchell, 2009; Blanton et al., 2007; Blanton et al., 2009). However, moving the field forward will require an agreement on how to settle outstanding questions. In this article, we analyze some of the evidence addressed in the target articles. We also review what implicit attitudes measures capture and describe possible avenues for empirical resolution of this debate.

The Definition of the Object Being Measured

The psychology of attitudes and attitude measurement has become a prominent and intellectually diverse field since its emergence in the 1920s (McGuire, 1986; Bolton & Danziger, 1991). One of the difficulties of measuring attitudes, however, is that people are generally unwilling to report attitudes that are socially undesirable. In fact, obtaining accurate measures of attitudes such as racial prejudice through explicit self-report is typically unsatisfying. In addition to people concealing some of their attitudes, they are often not fully aware of others and can construct socially desirable attitudes anew based on information from memory or the external environment (Schwarz, 2007; Schwarz & Bohner, 2007). Consequently, the challenge of how to measure socially undesirable attitudes such as racial prejudice remains.

One important milestone in implicit cognition research was Greenwald, McGhee, and Schwartz's (1998) development and validation of the Implicit Association Test (IAT), which quickly became the most widely used instrument to measure the newly defined implicit attitude. At the time, Greenwald and his colleagues stated that the IAT could gauge individual differences in people's positive and negative associations with an object even though people may not be aware of these associations. Although many scholars gladly accepted this notion, others were surprised by the IAT results and began to question what the instrument captures.

We wholeheartedly agree with Gawronski, Brownstein, and Madva on the need to understand whether the IAT captures hidden attitudes. With the advent of implicit attitude training to the workplace, this understanding is not only important to the field of psychology but also to society. Gawronski, Brownstein, and Madva's definition of implicit measures as behavior is unquestionably correct and helps to examine the IAT's validity. We concur that test-retest reliability is likely to be low when a measure reflects variable contextual cues, in addition to

being a function of individual, trait-like factors. We also agree that predictive validity is likely to be low if implicit measures vary across contexts.

Implicit attitude measures are best described as reflecting the temporary activation of knowledge stored in memory. This knowledge is largely related to the cultural milieu and social environments people inhabit, a point made early after the development of the IAT (Fazio & Olson, 2003; Olson & Fazio, 2004; Dambrun et al., 2008). In addition to the cultural bases of associations in memory, knowledge about social groups is related to personal experiences, and the activation of any representation varies with context (Rosseel et al., 2019; Olson & Fazio, 2004). Critical to the interpretation of implicit measures, the knowledge people recruit to respond to the IAT has evaluative implications because the measures present unambiguous evaluative concepts such as "good" and "bad." However, the knowledge activated is not identical to a summary attitude that categorizes the object or social group as "good" or "bad." Rather, implicit measures are like beliefs, affective reactions, and evaluations about *different* facets of an object. Classic models of attitudes and beliefs have shown that attitudes can be described as a summary evaluation of different beliefs, experiences, and feelings (Albarracin, 2021; Eagly & Chaiken, 1993). A group may be perceived to be aggressive, loud, and musical, and the evaluation of each these attributes will ultimately determine the summary attitude toward the group. However, the beliefs and evaluations of the attributes are not themselves summary attitudes[i], nor do they always contribute to summary attitudes in the same way. The complexity of these representations is likely to lead to less reliable implicit measures than one would obtain for summary attitudes.

The Evidence about Test-Retest Reliability, Predictive Validity, and Casual Efficacy

We appreciate that Machery provided a thorough examination of the reliability, validity, predictive power, and causal efficacy of implicit measures. With respect to reliability, Gawronski

and colleagues are correct that different contexts and times should bring up different associations and that the correlation among these associations cannot be expected to conform to the patterns of personality traits. In fact, the attitude research tradition departs from the personality scholarship in that attitudes are typically assessed for internal consistency at a single time point because many attitudes fluctuate along with social and historical changes. Thus, if summary measures of an object change over time, the complex associations gauged in implicit attitude measures should change even more.

We concur with Gawronski and colleagues that, for the same reasons, the predictive validity of implicit measures may be low when implicit measures and behaviors are obtained in different sessions. However, this area deserves scrutiny because most criterion measures are obtained in the same session as the implicit measures. In Table 1, we summarize the characteristics and main findings from all major meta-analyses of the correlations between implicit measures and behaviors. Across four meta-analyses, the correlation between implicit measures and behavioral criteria is modest (*r* ranging from .10 to .28) and lower than the associations between explicit attitudes and behaviors, (*r* ranging from.24 to .54; Aizen, Fishbein, Lohmann, & Albarracín, 2019, and average $r = .67$ for attitudes toward the behavior; Albarracín et al., 2001).

The meta-analyses of the implicit attitude-behavior correlation have important limitation. Specifically, the behavioral outcome is often loosely defined and sometimes departs from the norm in examinations of the correlation between explicit attitudes and behaviors (Glasman & Albarracín, 2006; Albarracín et al., 2001). In particular, both Greenwald et al. (2009) and Cameron et al. (2012) included behavioral intentions and judgments as “behavioral outcomes,” whereas meta-analyses of explicit attitudes and behaviors have required overt behaviors such as

purchasing a product, initiating an interaction, or performing a particular health behavior (see e.g., Glasman & Albarracín, 2006). Although Oswald et al. (2013) and Kurdi et al. (2019) reported including only behaviors, neither provided an operational definition of what behaviors were included.

If one is to take Cameron et al.'s (2012) meta-analysis of evaluative priming as an illustration, among a subset of 11 included reports that tested within-studies moderators, only five contributed overt behaviors, namely consumption of food, prosocial behavior, rejection of proposals in ultimatum games, drinking behavior, and behavior in a cyber ball game. The correlations between implicit measures and these behaviors ranged from .17 to .71, with the highest correlation involving attitudes towards alcoholic beverages. The strength of this specific correlation is notable and points to the possibility that affective associations with alcoholic beverages indeed drive behavioral choices. A likely reason is that consumers are likely to choose beer based on taste, temperature, and texture, whereas the decision to hire a white or Black candidate for a job is undoubtedly more complex. However, the variability of these correlations does not seem to be a function of measurement time because the behaviors in question were assessed in the same session as the implicit measures.

Another meta-analysis examined the success of procedures to change implicit measures by considering change in attitudes and behaviors (Forscher et al., 2019). In this synthesis, which did include measures of overt behavior, implicit bias training successfully changed implicit attitudes but had no overall impact on behaviors. These findings suggest that changes in implicit measures do not necessarily translate into changes of overt behaviors, and thus provide an additional layer of evidence that, in their current form, implicit measures are weakly correlated with overt behavior.

One problem with implicit measures is that neither a strong nor a weak association can conclusively support their validity and utility. In a case of "dammed if you do, dammed if you don't," the ideal result is neither a strong nor a weak correlation. In fact, the best evidence to support the utility of these measures would be for implicit attitudes to predict behavior over and above explicit attitudes when people try to conceal their real attitudes but for the implicit-attitude-behavior correlation to be as strong as the explicit-attitude-behavior correlation when behavior is neutral or socially desirable. The meta-analysis by Cameron et al. (2012), although important, has insufficient data to fully test this dissociation for overt behavior, and primary studies examining this dissociation have been criticized in the past (McConnell & Leibold, 2001; Blanton et al., 2009). However, their synthesis does have enough data to conclude that implicit attitudes are more strongly correlated with either judgments or behaviors in situations in which people are unlikely to control those judgments or behaviors.

All in all, implicit measures seem to be excellent instruments to understand activation of evaluatively relevant associations in specific contexts. However, using implicit measures to tap socially undesirable attitudes and predict overt behaviors will require further evidence of strong implicit-attitude-behavior correspondence in contexts where explicit attitudes do not predict the behavior. Meta-analyses and future primary studies should include measures of overt behaviors rather than rely on behavioral intention, trait judgments, or reports of anxiety (e.g., in Cameron et al., 2012).

How to Design Implicit Measures that Might Predict Behavior

Measures of implicit attitudes that predict overt behavior may benefit from applying the same methodological principles that were useful to resolve the attitude-behavior correspondence debate decades ago. For attitudes to predict behavior, the two measures must involve the same

action, target, time, and context (Ajzen & Fishbein, 1977; Ajzen & Fishbein, 1980). The measures should also be at the same level of specificity, such that a general attitude will predict a collection of behaviors more than a specific one (Fishbein & Ajzen, 1974). An additional complication for implicit attitudes is whether the names or pictures introduced in the implicit measure tap the beliefs or affective reactions that are activated at the time a behavioral decision is made. This problem is not simple, as the presentations in the test may not always match internal representations. To address a similar problem, Ajzen and Fishbein (1980) designed clear-cut procedures to choose what beliefs underlie an explicit attitude in a specific context. Similar methods could be explored for implicit attitudes.

In sum, attitude measures, either explicit or implicit, should predict behavioral outcomes more strongly when the measures correspond in level of specificity, time, target, action, and context. To enhance correspondence, implicit measures of attitudes should specify a behavior (e.g., refusing to help a Black homeless person) instead of a general target (e.g., Black people). In addition, if discrimination is more likely in private than in public (Gaertner & Dovidio, 1977; Saucier et al., 2005), this contextual factor should be integrated into implicit bias measures (Meissner et al., 2019). As one example, Kornadt et al.(2016) used the IAT to assess age stereotypes in different life domains (i.e., family domain and health domain) and for different age groups (i.e., younger, middle-aged, and older adults) and found distinct patterns across domains, highlighting the importance of contextual factors in implicit measures.

Developing other measures to predict socially undesirable behaviors may be useful as well. One possible direction would be to assess the attitude toward socially undesirable *in*actions instead of actions. Compared to actions, inactions are often seen as less intentional (Rosset, 2008; Sunderrajan & Albarracín, 2021), appear less consequential (Baron & Ritov, 2004),

receive less attention (Kahneman & Miller, 1986), elicit weaker emotional reactions (Landman, 1987; Zhou et al., 2010), and are judged as less negative even in the presence of undesirable outcomes (Sunderrajan & Albarracín, 2021). Therefore, people are likely to be more concerned with controlling their disposition toward socially undesirable actions than inactions. For example, the action of ridiculing a member from a minority group at a party should be judged more negatively and appear more consequential than the inaction of not approaching a member of the group at the same party. Therefore, people might be more willing to report their attitude toward inactions more accurately, and these attitudes may predict those inactions as well as active behaviors such as discrimination.

Final Note

The recent commentaries by Brownstein, Madva, and Gawronski and by Machery lead us to recommend increasing research attention to the correlation between implicit attitudes and overt behavior. In their present form, implicit measures are excellent instruments to gauge evaluatively relevant associations in specific contexts. Implicit attitudes also influence judgments of objects and persons when people fail to control the influence of their implicit attitudes (Cameron et al., 2012). However, past research on the association between implicit attitudes and overt behaviors has not clearly ascertained whether this association is stronger when people's ability and motivation to control the influence of their attitudes is lower than when it is higher. By improving implicit measures by measuring action, target, level, and context and ensuring high measurement correspondence with behavioral measures, future research should be able to settle this debate.

References

Aizen, I., Fishbein, M., Lohmann, S., & Albarracín, D. (2019). The influence of attitudes on behaviors. Albarracín, D., & Johnson, B.T. (Eds.). In D. Albarracin & B. T. Johnson (Eds.), *Handbook of Attitudes: Basic Principles, Second Edition.* (Vol. 1). Routledge.

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

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior* (Vol. 278). https://doi.org/Z

Albarracin, D. (2021). *Action and inaction in a social world: Predicting and changing attitudes and behaviors.* Cambridge University Press.

Albarracín, D., Johnson, B. T., Fishbein, M., & Muellerleile, P. A. (2001). Theories of reasoned action and planned behavior as models of condom use: a meta-analysis. *Psychological Bulletin*, *127*(1), 142.

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "would Jesse Jackson 'fail' the implicit association test?" *Psychological Inquiry*. https://doi.org/10.1207/s15327965pli1504_01

Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*. https://doi.org/10.1016/j.obhdp.2004.03.003

Blanton, H., Jaccard, J., Christie, C., & Gonzales, P. M. (2007). Plausible assumptions, questionable assumptions and post hoc rationalizations: Will the real IAT, please stand up?

In *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2006.10.019

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT. *Journal of Applied Psychology*. https://doi.org/10.1037/a0014665

Bolton, N., & Danziger, K. (1991). Constructing the Subject: Historical Origins of Psychological Research. *British Journal of Educational Studies*. https://doi.org/10.2307/3121154

Brownstein, M., Madva, A., & Gawronski, B. (2019). What do implicit measures measure? In *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1501

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.

Dambrun, M., Villate, M., & Richetin, J. (2008). Implicit racial attitudes and their relationships with explicit personal and cultural beliefs: What personalized and traditional iats measure. *Current Research in Social Psychology*.

Eagly, A. H., & Chaiken, S. (1993). Psychology of Attitudes. In *Psychology of Attitudes*.

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. In *Annual Review of Psychology*. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*(1), 59–74. https://doi.org/10.1037/h0035872

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B.

A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000160

Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of White racism, arousal, and helping behavior. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.35.10.691

Gawronski, B., Brownstein, M., & Madva, A. (2022). *How Should We Think About Implicit Measures and Their Empirical "Anomalies"?*

Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. Psychological Bulletin, 132(5), 778–822. https://doi.org/10.1037/0033-2909.132.5.778ming attitudes that predi. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.132.5.778

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*. https://doi.org/10.1037/0022-3514.74.6.1464

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

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*. https://doi.org/10.1037/0033-295X.93.2.136

Kornadt, A. E., Meissner, F., & Rothermund, K. (2016). Implicit and Explicit Age Stereotypes

for Specific Life Domains Across the Life Span: Distinct Patterns and Age Group Differences. *Experimental Aging Research*. https://doi.org/10.1080/0361073X.2016.1132899

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the implicit association test and intergroup behavior: A meta-analysis. *American Psychologist*. https://doi.org/10.1037/amp0000364

Landman, J. (1987). Regret and elation following action and inaction: Affective responses to positive versus negative outcomes. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167287134009

Machery, E. (2021). Anomalies in implicit attitudes research. In *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1569

Machery, E. (2022). *Anomalies in Implicit Attitudes Research: Not So Easily Dismissed*.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*. https://doi.org/10.1006/jesp.2000.1470

McGuire, W. J. (1986). The vicissitudes of attitudes and similar representational constructs in twentieth century psychology. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2420160202

Meissner, F., Grigutsch, L. A., Koranyi, N., Müller, F., & Rothermund, K. (2019). Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and

Sophisticated Solutions. In *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2019.02483

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653–667.

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.

Rosseel, L., Speelman, D., & Geeraerts, D. (2019). Measuring language attitudes in context: Exploring the potential of the Personalized Implicit Association Test. *Language in Society*. https://doi.org/10.1017/S0047404519000198

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*. https://doi.org/10.1016/j.cognition.2008.07.001

Saucier, D. A., Miller, C. T., & Doucet, N. (2005). Differences in helping whites and blacks: A meta-analysis. *Personality and Social Psychology Review*. https://doi.org/10.1207/s15327957pspr0901_1

Schimmack, U. (2021). Invalid Claims About the Validity of Implicit Association Tests by Prisoners of the Implicit Social-Cognition Paradigm. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691621991860

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*. https://doi.org/10.1521/soco.2007.25.5.638

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. *Blackwell Handbook of Social Psychology: Intraindividual Processes, 1*, 436–457.

Sunderrajan, A., & Albarracín, D. (2021). Are actions better than inactions? Positivity, outcome, and intentionality biases in judgments of action and inaction. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2021.104105

Tetlock, P. E., & Mitchell, G. (2009). Implicit Bias and Accountability Systems: What Must Organizations Do to Prevent Discrimination? In *Research in Organizational Behavior*. https://doi.org/10.1016/j.riob.2009.10.002

Zhou, Z., Yu, R., & Zhou, X. (2010). To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2010.08.010

Table 1
Summary of Existing Meta-Analyses of Correlations between Implicit Attitudes and Behavior

Short Reference	*Type of Implicit Measure*	*Type of Outcome*	*r*	*k*
Greenwald et al. (2009)	IAT	Behavior, reported behavior, judgment, physiological measures	$r = 0.274$	184
Cameron et al. (2012)	Sequential Priming	Behavior, intention, judgement	$r = 0.28$	167
Oswald et al. (2013)	IAT	Outcomes that arguably measured some form of discrimination	$r = 0.14$	298
Kurdi et al., (2019)	IAT	Behavior towards a group or members from a group	$r = 0.10$	217

r = Pearson's correlation involving implicit attitudes and behavioral criteria;
k = Number of reports included

[i] In philosophy, a belief is often understood as a propositional attitude that some proposition about the world is true. But in this article, the attitudes we refer to describe much broader and more complex attitudes that summarize all the beliefs, experiences and feelings, which is in line with the definition of attitudes in social psychology.