

Weak Criticisms and Selective Evidence: Reply to Blanton et al. (2009)

Allen R. McConnell
Miami University

Jill M. Leibold
Litigation Insights

We respond to a critique by H. Blanton et al. (2009), challenging our previous work demonstrating that an Implicit Association Test designed to assess implicit prejudice reliably predicts intergroup discrimination (A. R. McConnell & J. M. Leibold, 2001). We outline 3 flawed aspects of the critique. First, we note that claims that an outlier should be eliminated from the original data set are unfounded, and even with the elimination of this outlier, the conclusions of our original work are still strongly supported by the data. Second, we explain that concerns about interjudge reliability are specious and that considerable data support the validity of the judges' observations reported in our original study. Third, we note that claims of a disconnect between Implicit Association Test scores (argued to show negative bias against Blacks) and behavioral measures (argued to show pro-Black bias) are inappropriate because they neglect the relativistic nature of the key measures. Implications for the relation between implicit attitudes and behavior, for the law, and for future work on implicit attitudes are discussed.

Keywords: implicit attitudes, prejudice, discrimination, Implicit Association Test

Preface

We are pleased to offer our thoughts on the Blanton et al. (2009) article. In July 2006, we were contacted by an author on that article (G. Mitchell, personal communication, July 25, 2006), who asked us for copies of the data sets for our publication, McConnell and Leibold (2001). Despite the fact that the request came 5 years after publication and 8 years after data collection, we were happy to comply. At that time, we inquired about the nature of the interest in our data (e.g., what hypotheses were being tested), but no specifics were offered. In February 2007, we received a summary of their analyses, and we informed them that it is our practice to not offer comments on work outside of the normal editorial review process. We were not provided with a copy of their manuscript until June 2008, when we were asked to write a commentary on it after it had been peer reviewed. Thus, this reply is our first opportunity to speak to a number of reservations we have about their work. Herein, we focus on three concerns that were especially striking to us, and we elaborate on the broader implications of this exchange.

Selective Focus on a Single Dependent Measure

One of the points made by Blanton et al. (2009) is that it is tenuous to make strong claims about relations between implicit attitudes and behavior based on a small amount of data. We agree that it would be preferable to see more studies demonstrate the implications of implicit attitudes for behavioral relations and,

indeed, this was a primary motive behind our conducting the McConnell and Leibold (2001) study in the first place. One of the primary arguments marshaled by Blanton et al. against our work is that, in their view, our critical Implicit Association Test (IAT)–behavior demonstration hinges on 1 outlier participant, who revealed strong racial prejudice on the IAT (524-ms IAT effect, compared to the sample mean of 158 ms). Yet, it is interesting that this individual did not reveal the largest IAT effect in the entire sample (that distinction belonged to a 20-year-old woman at 578 ms). Thus, the choice of this critical outlier on statistical grounds seems questionable given that this person did not have the most extreme IAT score (a point not mentioned by Blanton et al.). Moreover, if a particular individual had an aberrant IAT score, this should make it more difficult to observe significant IAT–behavior correlations in our data set, not easier (because correlations require correspondence between two variables, and thus a deviant score on one measure would make such a relation less likely).

If this outlier did not have the most extreme IAT effect score in our sample, why do Blanton et al. (2009) claim that this individual should be discarded from the analyses? They argued her age (50 years) was problematic. We must admit that this is the first time we have encountered reservations about having a sample not composed exclusively of undergraduates aged 18–22 (there were also 2 participants in their 30s who, apparently, did not concern Blanton et al.). If one wishes to argue that nontraditional participants should be excluded from research, one must have a very compelling rationale for such a position. Blanton et al.'s argument is that older participants would have cognitive deficits (no citation or evidence is provided) and that research has shown that cognitive deficits can produce confounds on the IAT (McFarland & Crouch, 2002).

There certainly is considerable evidence that cognitive functioning declines with age (e.g., Foster & Taylor, 1920; Salthouse, Atkinson, & Berish, 2003), but the data also reveal that declines in normal individuals are relatively modest until around age 60, when the impact of aging becomes more pronounced (e.g., Hasher,

Allen R. McConnell, Department of Psychology, Miami University; Jill M. Leibold, Litigation Insights, Los Angeles, California.

This work was supported by National Science Foundation Grant BCS 0601148 and by the Lewis Endowed Professorship.

Correspondence concerning this article should be addressed to Allen R. McConnell, Department of Psychology, Miami University, Oxford, OH 45056. E-mail: mcconnar@muohio.edu

Stoltzfus, Zacks, & Rypma, 1991; Salthouse & Ferrer-Caja, 2003; Schroeder & Salthouse, 2004). Thus, the current critical outlier, in the absence of other evidence, is not a candidate for revealing strong cognitive deficits. Yet, Blanton et al. (2009) assert that the literature supports the position that such an individual would reveal inappropriate IAT scores. In particular, they cite McFarland and Crouch (2002), who concluded that general cognitive ability can introduce systematic method variance in the IAT, though those authors did not examine older adults in their work. But the Blanton et al. argument would require that a 50-year-old woman reveal sizable cognitive deficits, and such a premise does not comport with the existent literature (see Salthouse et al., 2003). Blanton et al. also cite an IAT study by Hummert, Garstka, O'Brien, Greenwald, and Mellott (2002) that examined older adults (mean age in their sample was 69 years) and their age-related attitudes using the IAT. Although Hummert et al. concluded that one should take general speed of responding (in the form of standardizing response latencies) into consideration when interpreting IAT results with older adults, those authors did not question the validity or utility of the IAT with such populations. Further, it is important to note that their average older participant was 19 years older than the outlier in question. Thus, it is unclear how one can argue that excluding this 50-year-old woman from the McConnell and Leibold (2001) analyses is justified. Moreover, because implicit attitudes reflect recurring paired associations between attitude objects and evaluations (e.g., Fazio, Eiser, & Shook, 2004; Jellison, McConnell, & Gabriel, 2004; Rydell & McConnell, 2006; Smith & DeCoster, 2000), it is quite probable that older individuals would, all things being equal, naturally have relatively stronger implicit prejudice against Blacks because they are more likely to have encountered Black–evaluation associative pairings during decades when cultural attitudes were even more negative toward Blacks. Consistent with this reasoning, Nosek et al. (2007) found that older individuals have the strongest racial IAT biases. Thus, rather than seeing this outlier with a relatively large bias against Blacks on the IAT as an anomaly (again, she did not have the most extreme IAT score in the sample), this logical outcome is what one would anticipate a priori.

Perhaps even more disconcerting than the questionable logic behind the argument to delete this individual from our analyses is that, upon closer examination of the data, the elimination of this participant does not really call into question the IAT–behavior finding in McConnell and Leibold (2001). It is curious that when challenging our findings, Blanton et al. (2009) focus exclusively on one single dependent measure, the judges' molar ratings of interactions, from a pool of several measures reported in our original article that were significantly correlated with the IAT. As a brief reminder to the reader, we assessed participant behavior toward both the Black and White experimenters (separately) using a number of converging indexes borrowed from the literature (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995). Specifically, the experimenters (both Black and White) evaluated the behavior of each participant using five different scales (e.g., participant friendliness, abruptness), which were combined into an overall index of experimenter perceptions. Also, trained judges watched videotaped interactions between the participant and each experimenter (Black and White) and for each made both global (molar) ratings (i.e., participant curtness, friendliness, and general comfort), and they also assessed a number of very specific behaviors associated with more positive interpersonal conduct (e.g., amount of smiling, speech errors made in interactions). Thus, there were three classes of behavioral variables (each based on difference scores between interactions involving the Black and White experimenters, such that larger scores reflected relatively more positive interactions with the White experimenter in comparison to the Black experimenter): experimenter ratings, judges' molar ratings, and judges' ratings of a number of specific behaviors.

Interestingly, Blanton et al. (2009) only focus on the correlation between implicit prejudice (as assessed by the IAT) and the judges' molar ratings and not on any of the other measures of bias that were significant in our original article. These other measures were all discussed extensively by McConnell and Leibold (2001), available to Blanton et al. in the data set provided to them, and were used as measures of biased interracial interactions in past published work. In Table 1, we provide a complete listing of how

Table 1
Zero-Order Correlations Between Implicit Association Test and Measures of Bias in McConnell and Leibold (None Eliminated) and Blanton et al. (Outlier Eliminated)

Measure of behavioral bias	Correlation conditions	
	None eliminated ($N = 41$; McConnell & Leibold, 2001)	Outlier eliminated ($N = 40$; Blanton et al., 2009)
Explicit measure of prejudice	.42**	.42**
Experimenters' ratings	.39*	.41*
Judges' molar ratings	.34*	.27[†]
Participant speaking time	.51**	.45**
Smiling	.39*	.40*
Speech errors	.42*	.36*
Speech hesitations	.35*	.25
Extemporaneous social comments	.31*	.35*

Note. All measures were coded such that larger, positive values reflect relatively more positive attitudes and behaviors toward Whites in comparison to Blacks. The boldfaced row is the only measure discussed by Blanton et al. (2009) in their critique of McConnell and Leibold (2001).

* $p < .05$. ** $p < .01$. [†] $p < .10$.

the IAT correlated to each and every behavioral measure (and also with explicit racial attitudes) that revealed a significant correlation in McConnell and Leibold (2001) as a function of whether all participants were included (left column; as in McConnell & Leibold) or whether the outlier participant was omitted (right column; as per Blanton et al.).¹

A few things should be noted. First, Blanton et al. (2009) focus on a different IAT index than the one we presented in our original publication. Specifically, McConnell and Leibold (2001) reported an IATFULL index that compared the average response latency for Blocks 3 and 4 to the average response latency for Blocks 6 and 7. Yet in their reanalysis, Blanton et al. focused on a different IAT index (IATRIM) that ignored Blocks 3 and 6, and thus relied on a comparison between Blocks 4 and 7. Although the differences between these two IAT indexes are small (e.g., the correlation between judges' molar ratings and the IATFULL is $r = .34, p < .05$, and the correlation between the judges' molar ratings and the IATRIM is $r = .32, p < .05$), it is unclear why Blanton et al. would emphasize a different measure (i.e., the IATRIM) than the one reported in the original article that they are critiquing. Thus, in the current article (see Table 1), we focus on the same IAT index reported in McConnell and Leibold.

As Table 1 reveals, when we eliminate the outlier, the once significant correlation between judges' molar ratings and the IATFULL measure drops from $r = .34, p < .05$, to $r = .27, p < .10$. In our judgment, the elimination of this outlier bumped this particular correlation a bit over the $p < .05$ barrier, but we do not view this change as monumental. But more important, Table 1 reveals that if Blanton et al. (2009) had reported on the predictive utility of the IAT for the other indexes of bias reported by McConnell and Leibold (2001), the vast majority of these correlations remained strong and significant even after we eliminated the outlier. Indeed, as people revealed relatively more negativity toward Blacks on the IAT, they had relatively more negative explicit attitudes toward Blacks; the White experimenters perceived relatively more positive interactions with the participants than did the Black experimenters; and the judges viewed that the participants showed less speaking time, less smiling, more speech errors, and fewer extemporaneous social comments toward the Black experimenter than the White experimenter. Each and every one of these findings is directly at odds with the Blanton et al. conclusion that the IAT did not predict behavioral bias in interracial interactions once the outlier was dismissed. Thus, out of the seven significant correlations between the IAT and biased behavior (as assessed both by the experimenters and by the judges) reported in McConnell and Leibold, five remained significant at conventional levels following the elimination of the outlier. Moreover, three of the correlations were larger following the elimination of the outlier too. Thus, we see considerable evidence that the IAT continues to predict indicators of biased intergroup behavior even with the elimination of this outlier.

It is curious why Blanton et al. (2009) only focus on one outcome measure, especially when it was one that produced the most relatively congenial results for their claims. Blanton et al. argued that they used the judges' ratings (instead of the experimenters' ratings) because "these are not filtered by the experience of actually participating in the interactions" (p. XX). This is a peculiar assertion for a number of reasons. First, Blanton et al. do not report the other ratings provided by the same judges (e.g., bias

in speaking time, smiling) that, based on the analyses reported in Table 1, do not support their conclusions. Second, work in the literature (e.g., Fazio et al., 1995) has relied on experimenters' experiences in interracial interactions as the key behavioral datum. In fact, who is better qualified to know whether an interaction was stilted than those involved? It is unclear how any bias could be introduced by the experimenters because they were completely unaware of each participant's prejudice level. Thus, the claim that their reports are inadmissible is hard to support. Indeed, one of the strengths of McConnell and Leibold (2001) was that we assessed both experimenters' experiences and independent judges' perceptions, and the strong correspondence between the experimenters' and judges' molar ratings ($r = .41, p < .01$) provides strong evidence for their validity.

In sum, the pattern of results observed, even with the elimination of the outlier, do not support the Blanton et al. (2009) conclusion that the correlation between the IAT and racially biased behavior was dependent on outliers. It is unclear why Blanton et al. do not note that these other correlations remained significant even when they (questionably) eliminated the outlier from their analyses. Further, it is also puzzling why Blanton et al. focus on a different IAT index than the one reported in the original article. It is understandable how a reader might conclude that such differences and omissions could reflect a selective presentation of the results, especially given that the (still) significant correlations that Blanton et al. fail to discuss are showcased in the original McConnell and Leibold (2001) article and that these data were in the data set provided to them.

Specious Concerns About Interjudge Reliability

Another point emphasized in the Blanton et al. (2009) critique is that interjudge reliability was problematic for the two judges who coded the videotapes. Despite suggestions to the contrary, the judges were extensively trained, and they spent 4 months painstakingly coding the videos. More important, in all cases, these judges showed significant correspondence across the 16 different dimensions they assessed for each of the two interactions between experimenters and the 41 participants. Although any study could benefit from greater interrater reliability, any dissimilarities between judges' ratings would only increase variability in their assessments, making it more (not less) difficult to observe the significant relations between the IAT and biased behaviors found in the study.

Moreover, if the judges were truly off base, their overall perceptions (i.e., judges' molar ratings) would have been discrepant with the experimenters' own reports, but there was reliable correspondence between them ($r = .41, p < .01$). Instead, the judges always showed significant agreement with each other, and in

¹ This exercise was the first time we had conducted analyses on these data in over 8 years. Our reanalysis replicated each and every finding reported in McConnell and Leibold (2001), with one exception. Specifically, the significant correlation between the IAT and bias in extemporaneous social comments shifted from .32 to .31 (both $ps < .05$). It is unclear what is responsible for this difference (e.g., rounding differences resulting from files being imported into sundry spreadsheet and statistical program packages, or differences in the software available today versus in 1999 when the original analyses were conducted).

concert their combined ratings significantly correlated with the independent reports of the experimenters. Thus, the general claim by Blanton et al. (2009) that the judges' ratings were unreliable is unsubstantiated. Consequently, their attempt to explore patterns of correlations for a particular judge seems unwarranted and post hoc.

Inappropriate Claims About a Disconnect Between IAT Scores and Behavioral Data

In their critique, Blanton et al. (2009) assert that there is a disconnect between participants' IAT scores (argued to be anti-Black) and behaviors (argued to be pro-Black). Specifically, Blanton et al. claim that because 37 of the 41 participants in the McConnell and Leibold (2001) study had positive IAT scores, these participants were biased against Blacks. However, viewing IAT = 0 as some sort of meaningful criterion is at odds with the understood relative nature of the IAT (see Greenwald, McGhee, & Schwartz, 1998; McConnell, Rydell, Strain, & Mackie, 2008; Nosek, Greenwald, & Banaji, 2006). What matters with the IAT is not an absolute score, but rather how one's relative standing on the measure predicts other outcomes, which themselves are relative. Also, Blanton et al. argue that participants in McConnell and Leibold were friendlier toward Blacks than toward Whites (based on examining judges' ratings for participants' interactions with the Black and White experimenters separately). Yet, such a characterization is misleading. For example, perhaps our Black experimenters were actually nicer people than our White experimenters and thus elicited more positive interactions. There is no way to equate, in an absolute sense, different people with different interactions, especially with subjective rating scales (see Biernat & Manis, 1994). The point is not whether one set of ratings (e.g., those for Black experimenters) is greater than another set of ratings (e.g., those for White experimenters). Instead, it is the relative difference between these sets of ratings (i.e., how friendly were participants toward members of one race relative to another) that is predicted by the IAT, which itself is a relativistic measure.

So What Is the Bottom Line of Blanton et al. Anyway?

In our last few sections, we address broader issues related to the Blanton et al. (2009) critique. First and foremost, the point of their criticism is lost upon us. They try to argue that implicit racial attitudes do not predict interracial behaviors. In their quest to establish this claim, they attack McConnell and Leibold (2001). And in the above sections, we explain why these criticisms are unfounded. However, even if one could successfully impugn McConnell and Leibold, other published studies show that implicit racial attitudes predict interracial behavior (e.g., Dovidio, Kawakami, & Gaertner, 2002; Fazio et al., 1995). That is, a number of diverse studies using a variety of measures to assess implicit racial attitudes come to the same bottom line (i.e., automatically activated attitudes predict more spontaneous forms of behavior). Thus, several studies support the same conclusions that McConnell and Leibold advanced, yet these well-established studies are not challenged (or even discussed) by Blanton et al.

Perhaps the intended message of the Blanton et al. (2009) critique is that a specific method (i.e., the IAT) is not a suitable candidate for predicting behavior, whereas other techniques (e.g., Dovidio et al., 1997; Fazio et al., 1995) are acceptable. If this is the

case, however, then other findings (also not discussed by Blanton et al.) in the literature demonstrating that the IAT predicts behavior are problematic for such a position. For instance, work such as Rydell and McConnell (2006) built implicit attitudes toward novel individuals and showed that implicit attitudes (as assessed by the IAT) but not explicit attitudes (as assessed by feeling thermometers and semantic differentials, just like McConnell & Leibold, 2001) predict spontaneous behaviors (e.g., greater seating distance from someone for whom a relatively negative bias was exhibited on an IAT). As a result, even if one questions McConnell and Leibold (2001), there is parallel evidence that the IAT itself can uniquely predict spontaneous behaviors. Thus, the Blanton et al. critique cannot successfully challenge the predictive utility of implicit attitudes in general or of the IAT in particular.

A Clarification on What We Mean by Behavior

One thing we must emphasize is that we have never claimed in our work (e.g., McConnell & Leibold, 2001; Rydell & McConnell, 2006; Rydell, McConnell, & Mackie, 2008) that implicit attitudes (as assessed by the IAT or by any other measure) should predict all forms of behaviors. In fact, we have underscored in past work that nonconscious associations (measured by a variety of approaches, including the IAT) are more likely to predict spontaneous behaviors than deliberative actions (see McConnell et al., 2008; Rydell & McConnell, 2006). Blanton et al. (2009) contend that the application of implicit attitudes (e.g., nonconscious racial prejudice) is problematic for the legal system. In our own work (including McConnell & Leibold, 2001), we have never claimed that implicit attitudes will predict effortful, deliberative courtroom judgments. In fact, that is why we have focused on measures of subtle behavior, such as nonverbal displays (e.g., McConnell & Leibold, 2001) and seating distance (e.g., Rydell & McConnell, 2006). Similar claims have been advanced by others (most notably Dovidio et al., 2002). That being said, we do believe that nonconscious biases (including implicit prejudice) can impact behavior and judgments in meaningful ways because like most dual-process systems (Chaiken & Trope, 1999), nonconscious processes typically have an automatic and unavoidable impact on perception and action that, later, more effortful processes may moderate. However, we anticipate that implicit prejudice would have a greater impact on interracial behaviors that are more spontaneous, less verbal, and less deliberative (see McConnell et al., 2008; Rydell & McConnell, 2006; Rydell, McConnell, Mackie, & Strain, 2006). However, the correspondence between implicit prejudice and relatively more deliberative interracial judgments and behaviors should increase considerably when one has reduced cognitive resources, an incorrect understanding of their own biases, or less motivation to render effortful decisions.

Implications for the Law

We contend that an appreciation of implicit attitudes and their implications for courtroom behavior is very important in understanding the outcomes of the legal system. Throughout a trial, jurors view evidence and testimony through the lens of their attitudes and beliefs. Processes such as *voir dire*, for example, are critical for identifying and removing prejudiced jurors. Yet during *voir dire*, jurors are asked to report their attitudes explicitly, which

can only reveal a subset of their feelings related to their behaviors and decisions (Dovidio et al., 2002; Rydell & McConnell, 2006; Wilson, 2002). Further, at this early stage in the legal process, jurors are prohibited from making any decisions and are often discouraged from expressing a bias, thus increasing the spontaneity and subtlety of their reactions to a particular case or party. Under idealized conditions, jury decisions should be quite effortful and thus less subject to the influences of nonconscious processes. Unfortunately, the courtroom is far from a perfect environment. For example, to the extent that it is a cognitively demanding place that taxes jurors' cognitive resources, implicit attitudes are more likely to predict guilt-related judgments (Strain & McConnell, 2008). Moreover, instructions to suppress thoughts (e.g., judge's instructions to ignore information or put attitudes aside) can, ironically, make those thoughts hyperaccessible because of rebound effects (Wegner, 1989), increasing the application of one's stereotypes and prejudices (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994). And more generally, jurors who are relatively unmotivated may just not engage in effortful judgment and decision making in the first place, such as in a recent mistrial in Sydney, Australia, where jurors hearing a case involving a possible life sentence played Sudoku instead of paying attention to the proceedings because they were bored (ABC News, 2008).

These insights are especially important because decades after the Civil Rights Act of 1964 and the Age Discrimination in Employment Act of 1967, sex, race, religion, and other employment discrimination claims are still being filed, with age and family responsibility discrimination cases, in particular, being filed in ever-increasing numbers (Hodge, 2004; Still, 2006; United States Equal Opportunity Employment Commission, 2007). In fact, recently the Supreme Court shifted the burden of proof onto employers in age discrimination cases to show that their actions toward an employee were nondiscriminatory and taken for reasons other than age (*Meacham v. Knolls Atomic Power Lab*, 2008). A burden shift such as this is a key reflection of the level of discrimination that the highest court sees in the United States today. Yet, these cases are notoriously difficult to prove because the discrimination occurs over time and often in subtle ways when employers have discretion in subjective ratings of employee performance (*Meacham v. Knolls Atomic Power Lab*, 2008). Employers may pay women less or lay off older workers, but they probably do not often explicitly rationalize these decisions as a belief that "it's a man's job," "women will not put in the time and effort it takes," or "older employees are ineffective" (Fiske, 1998; *Walker v. Fred Nesbit Distrib. Co.*, 2004). Thus, subtle and nonconscious biases can have important implications in the workplace as well as in the courts that hear worker discrimination cases. Although we would not contend that implicit attitudes and prejudices are the sole determinant of biased behaviors in the workplace and courtroom, we do believe they play a role, and in some circumstances a very consequential role, in determining critical outcomes. Because bias in the courtroom is especially insidious given the basic precepts of jurisprudence in our society, minimizing its effects among empanelled jurors is especially important. But sadly, it is well documented that jurors have difficulties in ignoring specific features of cases or in discounting their own biases (Sunstein, Hastie, Payne, Schkade, & Viscusi, 2003), especially in situations where their bias is nonconscious and thus jurors would not have a starting point from which to consider how to correct any biases they

possess (Wegener & Petty, 1997). In sum, it is clear that the relations between implicit prejudice and discriminatory behavior are nontrivial, and real implications exist for the legal system.

Implications for Future Progress in Implicit Attitudes and Prejudice

We were pleased that many years after the publication of McConnell and Leibold (2001), additional analyses (by us and by others) bore out consistent findings reaffirming that implicit attitudes, as assessed by the IAT, reliably predicts intergroup behaviors in ways that are detectable by both those involved in the interactions and by third-party observers alike. These findings are consistent with other intergroup relations studies (e.g., Dovidio et al., 2002; Fazio et al., 1995) and with others where implicit attitudes were built in the lab and assessed with the IAT (e.g., Rydell & McConnell, 2006). The current work demonstrates that despite the concerns expressed by Blanton et al. (2009), the conclusions offered by McConnell and Leibold are quite robust even when under considerable scrutiny.

That being said, we support those who challenge existent work, especially in the face of a relatively small literature on IAT-behavior relations. It was in this spirit that we freely provided all of our data upon request to Blanton et al. (2009) because we believe that doing so is important for providing a transparent account of one's work. Indeed, scientists have an obligation to conduct research with the utmost of integrity and respect for the scientific enterprise and to subject themselves to its scrutiny. Further, it is clear that Blanton et al. experienced difficulty in securing some relevant data sets, and we can understand how frustration can result from this. With this in mind, we offer some observations that might prove useful in encouraging the free exchange of ideas and data in the future.

We believe that trust is the cornerstone of scientific openness and that trust is a two-way street. First, those who wish to critique findings must be able to trust that data will be available to them in a timely and unfettered fashion. Not only is this consistent with American Psychological Association guidelines and procedures but, more importantly, by the recognition that data (and not reports of data) are the basic building blocks of our field. The passage of time may lead to reinterpretations of findings, but in the end, data are the foundation of researchers' theories and future work. However, scientists must know that fellow researchers will treat them fairly and respectfully. For example, when skeptics can articulate a priori hypotheses regarding their interest in others' data, everyone benefits. Those skeptics will have a guiding framework for their inquires, researchers whose data are reviewed will feel that the likelihood of a "fishing" or "cherry picking" expedition is minimal, and both parties can be maximally helpful to each other in identifying important facets of the work (e.g., auxiliary data, unreported observations) that can provide a more complete picture of the work conducted. And more important, strong conclusions cannot be drawn simply from reexaminations of data. Although such inquires may identify different interpretations for existent data, it is then incumbent upon researchers to conduct primary research to establish such claims conclusively (e.g., experimentally manipulate the hypothesized confounds). When considering the critique offered by Blanton et al. (2009), perhaps those who are older might reveal stronger IAT-behavior correspondence. But

rather than viewing this as a problem for the IAT, this insight can be parlayed into testable hypotheses with a strong theoretical foundation. For instance, to the extent that automatically activated attitudes are more likely to guide behavior when cognitive expenditures are reduced (e.g., Fazio & Towles-Schwen, 1999), perhaps those who are older are less likely to inhibit their deliberately chosen behaviors (Hasher & Zacks, 1988) and thus respond more in line with their implicit prejudices in intergroup interactions. Maybe such speculations are incorrect, but they are certainly testable. Importantly, such ideas can lead to new research that pushes theory on implicit prejudice forward in compelling and sophisticated ways.

Without trust, progress in any domain of science will be stifled. And to the extent that particular researchers on either side of an intellectual divide repeatedly arouse mistrust and suspicion, whether real or perceived, walls to an open and transparent scientific enterprise will be erected and reinforced. We hope that, in addition to defending the credibility of our past work, that the current episode illustrates the benefits of an open scientific exchange of ideas that can push researchers' understanding of important psychological phenomena and principles forward.

References

- ABC News. (2008, June 11). *Drug trial abandoned after jurors play Sudoku*. Retrieved January 2, 2009, from <http://www.abc.net.au/news/stories/2008/06/11/2270736.htm>
- Age Discrimination in Employment Act of 1967, Pub. L. No. 90–202, 81 Stat. 602.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*, 5–20.
- 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, 94*, XX–XX.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Civil Rights Act of 1964, Pub. L. No. 88–352, 78 Stat. 241.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82*, 62–68.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.
- Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*, 293–311.
- Fazio, R. H., Jackson, J. R., Dutton, 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*, 1013–1027.
- Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 97–116). New York: Guilford Press.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 357–411). New York: McGraw-Hill.
- Foster, J. C., & Taylor, G. A. (1920). The applicability of mental tests to persons over 50. *Journal of Applied Psychology, 4*, 39–58.
- 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*, 1464–1480.
- Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 163–169.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.
- Hodge, K. H. (2004). *Employment discrimination and civil rights actions in federal and state courts*. Philadelphia: American Law Institute, American Bar Association Continuing Legal Education.
- Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the Implicit Association Test to measure age differences in implicit social cognitions. *Psychology and Aging, 17*, 482–495.
- Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and related behaviors and beliefs among gay and straight men. *Personality and Social Psychology Bulletin, 30*, 629–642.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808–817.
- 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, 37*, 435–442.
- McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology, 94*, 792–807.
- McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20*, 483–510.
- Meacham v. Knolls Atomic Power Lab., 128 S. Ct. 2395 (2008).
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.
- Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Linder, N. M., Ranganath, K. A., et al. (2007). Pervasiveness and correlations of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*, 36–88.
- Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*, 995–1008.
- Rydell, R. J., McConnell, A. R., & Mackie, D. M. (2008). Consequences of discrepant explicit and implicit attitudes: Cognitive dissonance and increased information processing. *Journal of Experimental Social Psychology, 44*, 1526–1532.
- Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954–958.
- Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132*, 566–594.
- Salthouse, T. A., & Ferrer-Caja, E. (2003). What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychology and Aging, 18*, 91–110.
- Schroeder, D. H., & Salthouse, T. A. (2004). Age-related effects on cognition between 20 and 50 years of age. *Personality and Individual Differences, 36*, 393–404.
- Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4*, 108–131.

Still, M. C. (2006). *Litigating the maternal wall: U.S. lawsuits charging discrimination against workers with family responsibilities*. Retrieved July 3, 2008, from <http://www.worklifelaw.org/pubs/FRDreport.pdf>

Strain, L. M., & McConnell, A. R. (2008). *Happy to punish you: Mood moderates the relation between implicit attitudes and discriminatory judgments*. Unpublished manuscript.

Sunstein, C. R., Hastie, R., Payne, J. W., Schkade, D. A., & Viscusi, W. K. (2003). *Punitive damages: How juries decide*. Chicago: University of Chicago Press.

United States Equal Opportunity Employment Commission. (2007, February 1). *Enforcement statistics and litigation*. Retrieved July 4, 2008, from <http://www.eeoc.gov/stats/enforcement.html>

Walker v. Fred Nesbit Distrib. Co., 331 F. Supp. 2d 780 (S.D. Iowa 2004).

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Mahwah, NJ: Erlbaum.

Wegner, D. M. (1989). *White bears and other unwanted thoughts: Suppression, obsession, and the psychology of mental control*. New York: Viking/Penguin.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.

Received July 4, 2008

Revision received October 1, 2008

Accepted October 23, 2008 ■

ORDER FORM

Start my 2009 subscription to the *Journal of Applied Psychology* ISSN: 0021-9010

___ \$98.00 **APA MEMBER/AFFILIATE** _____

___ \$249.00 **INDIVIDUAL NONMEMBER** _____

___ \$715.00 **INSTITUTION** _____

In DC add 5.75% / In MD add 6% sales tax _____

TOTAL AMOUNT DUE \$ _____

Subscription orders must be prepaid. Subscriptions are on a calendar year basis only. Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN
PSYCHOLOGICAL
ASSOCIATION

SEND THIS ORDER FORM TO
American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

Call **800-374-2721** or 202-336-5600
Fax **202-336-5568** :TDD/TTY **202-336-6123**
For subscription information,
e-mail: **subscriptions@apa.org**

Check enclosed (make payable to APA)

Charge my: Visa MasterCard American Express

Cardholder Name _____

Card No. _____ Exp. Date _____

Signature (Required for Charge)

Billing Address

Street _____

City _____ State _____ Zip _____

Daytime Phone _____

E-mail _____

Mail To

Name _____

Address _____

City _____ State _____ Zip _____

APA Member # _____

APLA09