


RESEARCH NOTE

The BIAT and the AMP as measures of racial prejudice in political science: A methodological assessment

Katherine Clayton* , Jordan Horrillo and Paul M. Sniderman

Department of Political Science, Stanford University, Stanford, USA

*Corresponding author. Email: kpc14@stanford.edu

(Received 24 November 2021; revised 8 June 2022; accepted 13 September 2022; first published online 1 December 2022)

Abstract

Political scientists often use measures such as the Brief Implicit Association Test (BIAT) and the Affect Misattribution Procedure (AMP) to gauge hidden or subconscious racial prejudice. However, the validity of these measures has been contested. Using data from the 2008–2009 ANES panel study—the only study we are aware of in which a high-quality, nationally representative sample of respondents took both implicit tests—we show that: (1) although political scientists use the BIAT and the AMP to measure the same thing, the relationship between them is substantively indistinguishable from zero; (2) both measures classify an unlikely proportion of whites as more favorable toward Black Americans than white Americans; and (3) substantial numbers of whites that either measure classifies as free of prejudice openly endorse anti-Black stereotypes. These results have important implications for the use of implicit measures to study racial prejudice in political science.

Keywords: Affect Misattribution Procedure; Brief Implicit Association Test; implicit bias, racial prejudice

Understanding the role that racial prejudice plays in American politics is a major challenge for political scientists. Until recently, research relied primarily on explicit measures of racial prejudice. However, these measures are vulnerable to social desirability bias. As norms of egalitarianism have grown increasingly entrenched in American society, some prejudiced people have become less willing to openly express discriminatory attitudes against members of other racial or ethnic groups (Huddy and Feldman, 2009).

Facing this challenge, the introduction of measures of *implicit* prejudice promised a breakthrough. The Implicit Association Test (IAT) developed by Greenwald *et al.*, for example, was designed to reveal “race-related stereotypes and attitudes that are consciously disavowed by the subjects who display them” (Greenwald *et al.*, 1998, p. 1473). Similarly, the Affect Misattribution Procedure (AMP) claimed “to circumvent responses strategies . . . among individuals or in situations where motivational pressures are high” (Payne *et al.*, 2005, p. 289). The IAT, the BIAT (a shorter version of the IAT developed by Sriram and Greenwald (2009)), and the AMP were quickly embraced by political scientists as “alternative survey techniques for measuring racial prejudice in ways that circumvent social desirability pressures” (Segura and Valenzuela, 2010, p. 501).

Implicit measures of prejudice gained substantial traction in political science when Barack Obama ran for president in 2008, with scholars launching innovative studies designed to uncover the impact of hidden or subconscious “anti-African American racism” on voting behavior (Pasek *et al.*, 2009; see also Greenwald *et al.*, 2009; Finn and Glaser, 2010; Payne *et al.*, 2010; Segura and Valenzuela, 2010; Messing *et al.*, 2016). Subsequent research used implicit measures to

understand prejudice toward other marginalized groups including immigrants (e.g., Malhotra *et al.*, 2013), women (e.g., Mo, 2015), and Hispanics (e.g., Kam, 2007; Pérez, 2010, 2016). Across these studies, implicit measures were seen as “very useful in studying politically relevant attitudes that individuals will often want to misrepresent” (Malhotra *et al.*, 2013).

In social psychology, however, the inevitable cycle of critical assessment following innovation had begun. Early on, scholars waged that the IAT may capture stereotype *awareness* rather than endorsement (Arkes and Tetlock, 2004). Now, both proponents and critics of implicit measures in social psychology agree that scores on implicit prejudice measures should *not* be interpreted as capturing racial prejudice that is unaffected by social desirability pressures (e.g., Fazio and Olson, 2003; De Houwer *et al.*, 2007; Ito *et al.*, 2015; Gawronski *et al.*, 2017). Rather, implicit prejudice may reflect just one side of a dual-process model (e.g., Gawronski and Bodenhausen, 2006; Gregg *et al.*, 2006)—a separate cognitive system, with fundamentally different consequences for behavior than explicit prejudice (e.g., Dovidio *et al.*, 2002), especially in the context of sensitive topics like race and discrimination (e.g., Greenwald *et al.*, 2009). Many social psychologists, including the developers of the two measures of implicit prejudice most widely used in political science, the IAT/BIAT and the AMP, also agree that the concept of “implicit prejudice” should be discarded because it is systematically ambiguous and misleading (e.g., Greenwald and Banaji, 2017; Payne *et al.*, 2017; Corneille and Hütter, 2020).

For their part, political scientists have demonstrated that when explicit and implicit measures of prejudice are pitted against one another, explicit measures better predict outcomes of most interest to political scientists, such as vote choice or policy views (Ditonto *et al.*, 2013; Kalmoe and Piston, 2013; Kinder and Ryan, 2017). Nevertheless, studies of racial attitudes and their political effects—including those published in the discipline’s top journals—have continued to incorporate the IAT, the BIAT, or the AMP as prejudice measures on the grounds they that circumvent the underreporting of racial bias caused by “normative pressures facing respondents asked explicit questions about race relations” (Iyengar and Westwood, 2015, p. 696; see also Valentino *et al.*, 2018; Chudy, 2021; Engelhardt, 2021). The objective of this study is to assess the credibility of claims that the measures of implicit prejudice most commonly used by political scientists, the BIAT and the AMP, provide valid measures of hidden racial prejudice.

We take a different approach from prior work in three main ways. First, we conduct the first analysis of implicit measures of prejudice in a high-quality, nationally representative sample of respondents—the only such sample to our knowledge in which respondents completed both the BIAT and the AMP. Prior work in social psychology has examined the correlation between the IAT/BIAT and the AMP on small samples and/or anomalous samples of respondents coming forward to be interviewed because of their interest in racism (e.g., Payne *et al.*, 2008; Greenwald *et al.*, 2009; Bar-Anan and Nosek, 2014).¹ In contrast, our sample comes from the American National Election Study—the gold standard in political science for nationally representative surveys.

Second, we take a different approach from work that reports levels of implicit anti-Black prejudice by evaluating whether the measures overstate favorability toward Black people among white Americans. Finally, we build on previous political science research that compares the relative explanatory power of implicit and explicit measures to examine how well implicit measures of prejudice capture anti-Black sentiments among the most explicitly prejudiced white respondents in our samples.

Our results bring out key limitations of both the BIAT and the AMP as measures of hidden or unconscious racial prejudice. First, we demonstrate that the relationship between the Black/white BIAT and the Black/white AMP, although statistically significant, is substantively indistinguishable from zero. Given that political scientists have used both to measure the same

¹These studies generally find weak correlations between the measures, ranging from 0.11 (Payne *et al.*, 2008) to 0.24 (Greenwald *et al.*, 2009; Bar-Anan and Nosek, 2014).

construct—anti-Black prejudice—one or both measures cannot be valid. Second, we show that both the BIAT and the AMP classify one in every three white respondents as not merely free of racial prejudice, but as preferring Black people to white people. Decades of research in political science and public opinion have made clear that the claim that a third of white people are biased in favor of Black people is not credible. Third, our analyses bring to light that implicit racial prejudice measures frequently fail to identify even the most racially prejudiced whites. Substantial numbers of white Americans who explicitly declare that Black Americans are less intelligent and lazier than white Americans are classified as free of anti-Black prejudice by the BIAT and/or the AMP.

The paper proceeds as follows. First, we summarize how the BIAT and the AMP were administered in our data sources. Next, we present the main empirical results. Then, we highlight and evaluate a number of methodological concerns. Finally, we call attention to some broader implications of our findings.

1. Data and methods

The BIAT and the AMP were administered in the 2008–2009 ANES panel study. The AMP was also administered in the 2008 ANES time series study. Our analyses rely primarily on the panel, but we also make use of the time series for the purposes of comparison. Given that the methodology of the BIAT and the AMP as administered in the ANES is primarily designed to gauge anti-Black prejudice among white individuals, we restrict our analyses to white respondents only.

The ANES panel used a brief version of the IAT (BIAT), developed to shorten the time required to administer the test (Sriram and Greenwald, 2009). While the basic methodology of each test is nearly identical, we consider some concerns associated with the brief version in the discussion section. Both the BIAT and the IAT instruct respondents to press a keyboard key as quickly as they can after seeing one of four different kinds of text or visual stimuli on a screen (a Black person's face, a white person's face, a positive word, or a negative word) in a series of repeated trials. Specifically, they are instructed to press the same key for white faces and for negative words and another key for anything else, or the same key for Black faces and for positive words and another key for anything else. The next round alternates, so participants must classify white faces with the positive category and Black faces with the negative category. Based on the difference in response times between white-good, Black-bad, and white-bad, Black-good, a D-score is calculated on a scale of -2 to 2 , where -2 indicates maximum preference for Black people over white people and 2 maximum preference for white people over Black people.

In the AMP, respondents are first shown a picture of a Black person's face or a white person's face on a screen for a fraction of a second, followed by a picture of a Chinese character displayed for a longer time. They are then asked to say whether the Chinese character appeared pleasant or unpleasant to them. Crucially, they are reminded that the photographs they saw prior to the Chinese character might bias their answers and are specifically instructed to guard against this. The resulting AMP scores are calculated on a scale that ranges from -1 to 1 , where -1 indicates that respondents classify all characters preceded by a Black person's face as pleasant (maximum pro-Black preference) and those preceded by a white person's face as unpleasant, and 1 indicates the opposite.

The distinctive feature of our study is that the same respondents took both the BIAT and the AMP. The final sample size for our main analysis of white respondents who completed both measures of implicit prejudice is 1352. The AMP was administered online during Waves 9 and 10 of the panel.² The BIAT was administered in Wave 19. Following standard practice (Kinder and

²Two versions of the AMP were administered in a random order in the panel (one in each wave)—a version with non-famous Black and white faces, and an alternative version showing the faces of Barack Obama and John McCain (excluded from our analysis). Kalmoe and Piston (2013) find some evidence the Obama-McCain AMP in Wave 9 could have contaminated responses to the Black-white AMP in Wave 10. When we subset our main results to include only those respondents who took the Black-white AMP first ($N = 679$), however, they are substantively identical (see Figure B3 in the appendix).

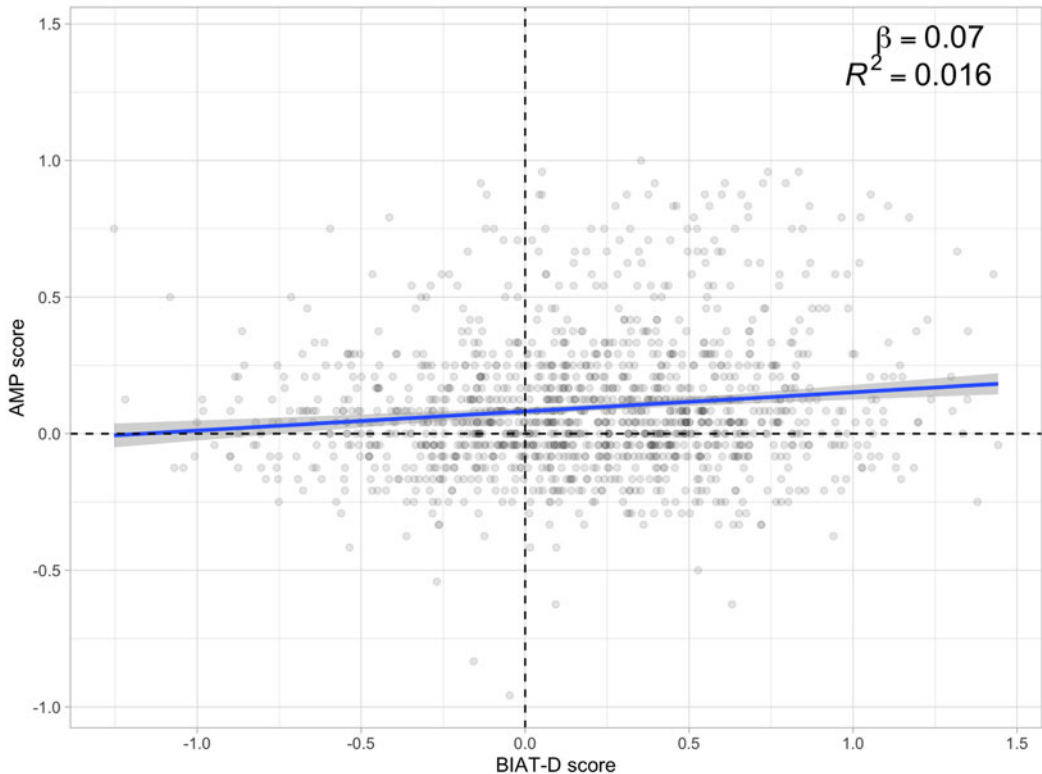


Figure 1. Relationship between BIAT-D scores and AMP scores is virtually non-existent. Note $N = 1352$ white respondents in the 2008–2009 ANES panel study. BIAT-D scores are measured on a -2 to 2 scale, and AMP scores on a -1 to 1 scale (higher values indicating higher anti-Black prejudice).

Ryan, 2017), we exclude respondents who evaded valid AMP measurement by selecting either “unpleasant” or “pleasant” after every profile they viewed, about 10 percent of respondents ($N = 158$), or who responded too rapidly, too slowly, or had an error rate above 35 percent on the BIAT (7 percent of the full sample; $N = 105$).³

2. Results

Political scientists routinely rely on the BIAT and the AMP as measures of covert racial prejudice that circumvent social desirability biases. If they both measure hidden racial prejudice, scores on one will predict scores on the other well. Figure 1 plots white respondents’ BIAT-D scores on the x-axis and their AMP scores on the y-axis ($N = 1352$) in the 2008–2009 ANES panel study. For each implicit measure, higher values indicate a positive preference for white people over Black people; conversely, negative values indicate a positive preference for Black people over white people, and zero indicates indifference. The solid line is an OLS regression line and the gray shaded area around the line is the 95 percent confidence interval. The quantity of interest is the magnitude of the relationship between the BIAT and the AMP.

As Figure 1 shows, there is virtually no connection between them ($b = 0.07$, $R^2 = 0.016$). True enough, a one-unit movement on the BIAT produces a 0.07-unit movement on the AMP and that

³In the 2008 ANES time series study, the AMP was included in the post-election questionnaire, which was administered in November–December 2008. After excluding 105 respondents who evaded valid AMP measurement and 37 respondents with missing outcome data, our final sample size for the time series study is 894 white respondents.

relationship is statistically significant at conventional levels—but the substantive effect is essentially zero. An individual’s score on the BIAT tells us virtually nothing about her score on the AMP, and her score on the AMP tells us virtually nothing about her score on the BIAT. This means that either the BIAT or the AMP is not capturing hidden racial prejudice, or both measures are not capturing hidden racial prejudice.

To examine whether either is a credible measure of covert racial prejudice, we look at the BIAT and the AMP individually. **Figure 2** shows density plots of white respondents’ BIAT-D scores (left) and AMP scores (middle) in the ANES panel and, for cross-validation, the results from the AMP administered in the 2008 ANES time series study (right).

Consistent with previous research (e.g., Greenwald and Banaji, 2017; Greenwald and Lai, 2020), the unshaded area in **Figure 2** to the right of the zero point of indifference (dashed line) shows that most white Americans “demonstrate automatic preference for whites relative to Blacks” (Greenwald and Lai, 2020, p. 426). This is the result that has provided a legal predicate for anti-discrimination suits and propelled the widespread use of measures of implicit prejudice in racial sensitivity training. The problem is that the same measurement procedure classifies implausibly large numbers of white Americans as having an automatic preference for Black people relative to white people. The shaded area to the left of the zero line in **Figure 2** shows the percentage of respondents who are classified as more favorable toward Black people than white people. For both the BIAT and the AMP in the panel study, it is 34 percent. For the AMP in the ANES time series study, it is 30 percent, indicating that respondents who took the AMP in the time series behaved similarly to those in the panel.

It is important to note that this result follows from the standard calculation of a zero point of indifference. Greenwald *et al.* (2006) demonstrate that the zero point on the IAT maps onto the zero point of self-reported, explicit measures of preference. Moreover, the calculation of a zero point of indifference continues to be standard operating procedure in research on implicit measures (see Greenwald and Lai, 2020). In fact, among respondents classified as more favorable to Black people than white people on the BIAT, the mean BIAT-D score is -0.31 (on a -2 to 2 scale). Among those classified as pro-Black on the AMP in both the panel and the time series, the mean AMP score is -0.13 (on a -1 to 1 scale). The ANES data thus indicate that apparent implicit preferences for Blacks over whites are fairly strong among many white respondents.

The claim that one out of every three white Americans has an automatic preference for Black Americans relative to whites is improbable, given decades of political science research on white racial attitudes (Huddy and Feldman, 2009 provide a useful review). In the 2008 ANES time series

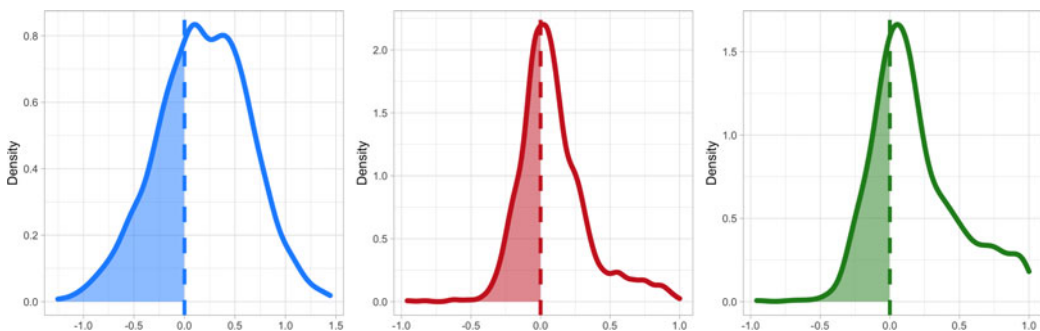


Figure 2. Both the BIAT and the AMP overstate white respondents’ favorability toward Black people. *Note* $N = 1352$ white respondents in the 2008–2009 ANES panel study; $N = 894$ white respondents in the 2008 ANES time series study. Each plot shows the density of scores on implicit measures of prejudice across all respondents. The dashed line is the zero line and the shaded area to the left of the line shows the fraction of whites who appear to be more favorable toward Black people than whites on each implicit measure.

Table 1. Significant proportions of white respondents who openly express prejudice toward Black people are classified as free of implicit anti-Black prejudice

| | Non-prejudiced, IAT (panel) | Non-prejudiced, AMP (panel) | Non-prejudiced, AMP (time series) |
|--|--------------------------------|--------------------------------|--------------------------------------|
| Rate Black people lazier than white people | 23% | 24% | 20% |
| Rate Black people less intelligent than white people | 26% | 28% | 19% |
| Feel cooler toward Black people than white people | 26% | 28% | 23% |

Explicitly prejudiced white respondents only are included. In the first row, $N = 228$ white respondents in the 2008–2009 ANES panel study and $N = 410$ in the time series study rate Black people as lazier than white people; in the second row, $N = 347$ in the panel and $N = 361$ in the time series rate Black people as less intelligent than white people; in the third row, $N = 267$ in the panel and $N = 297$ in the time series feel cooler toward Black people than white people. The values in each cell are the percentages of these explicitly prejudiced respondents who are coded as not prejudiced against Black people (either neutral or more favorable toward Black people than white people) based on their BIAT-D-score (second column) or AMP score (panel, third column; time series, fourth column).

study, for example, only 9.7 percent of white respondents rate their feelings toward Black people as warmer than their feelings toward white people on a standard feeling thermometer.⁴ Likewise, only 3.6 percent of white respondents rate Black people as harder working, and just 2.1 percent as more intelligent, than white people. It is not easy to square the finding that, according to the BIAT and the AMP, one in every three whites prefers Black people relative to white people with the claim that implicit measure of racial prejudice is a satisfactory method for circumventing social desirability bias.⁵

Our final analysis builds on previous work that has evaluated the relative explanatory power of implicit and explicit measures of prejudice (e.g., Ditonto *et al.*, 2013; Kalmoe and Piston, 2013; Kinder and Ryan, 2017). These studies usually include both implicit and explicit measures of prejudice in regression analyses to evaluate their effect on outcomes like policy preferences or vote choice, and find that the explanatory power of the implicit measures declines substantially when explicit measures are accounted for. We take a different approach, subsetting our ANES data to include just those respondents who are willing to openly endorse explicitly prejudiced statements.

Specifically, the first row of Table 1 under the column headers subsets our data to include respondents who rate Black people as “lazier” than white people. The second row includes those who rate Black people as “less intelligent at school” than white people (“less intelligent” in the time series). The third row comprises the those who feel “cooler” toward Black people than whites. As Table 1 shows, among those who describe Black people as lazier than whites, 23 percent are classified as free of anti-Black prejudice on the BIAT in the 2008–2009 panel, 24 percent are classified as free of prejudice on the AMP in the panel, and 20 percent are classified as free of prejudice on the AMP in the 2008 time series. The comparable numbers for white respondents who say Black people are less intelligent than whites are 26, 28, and 19 percent. Finally, among those who feel “cooler” toward Black people than whites, 26, 28, and 23 percent in the BIAT (panel), AMP (panel), and AMP (time series), respectively, are classified as free of implicit anti-Black prejudice. Again, it is not obvious how to square the claim that the measures

⁴We rely on the time series here because it employs the standard and most commonly used wording for both the feeling thermometer and group stereotype questions (see the appendix for details), thus making it most directly comparable to existing political science research on explicit prejudice. An alternative question format for explicit prejudice measures was used in the panel; we examine it in detail in Table I.

⁵There are explicit questions in the ANES panel study that show substantial numbers of whites sympathetic to Blacks (e.g., 33 percent of white respondents say Black people have too little influence in American politics and 42 percent say that discrimination holds Black people back). See Chudy (2021) for an innovative analysis of racial sympathy. In contrast, the focus here is measures of racial prejudice, and white respondents are generally unwilling to report favorability toward Black people at their expense (i.e., by expressing more positive affect toward Black people than fellow whites). Since the BIAT and the AMP compare reactions to Black and white faces, the stereotype and feeling thermometer questions that compare responses across groups are more appropriate explicit measures for the purposes of comparing implicit and explicit measures.

of implicit prejudice used in political science circumvent social desirability biases when they classify between a fifth and third of whites who are willing to tell a stranger, the ANES interviewer, that Black people are inferior to white people as free of anti-Black racial prejudice.

3. Discussion

Many political scientists use the BIAT and/or the AMP in research on racial prejudice because they claim that the measures circumvent social desirability bias, providing a clear picture of hidden or unconscious racism. However, the main justification behind using these measures to overcome socially desirable responding has consisted primarily in descriptions of their procedures, not in the presentation of evidence that they in fact do so. This is the first study based a high-quality, nationally representative sample to critically examine the validity of the BIAT and the AMP. Our results call into question the claim that either implicit test provides a valid measure of covert prejudice. This is all the more reason to consider whether our results are a product of methodological choices of the 2008–2009 ANES panel or the 2008 ANES time series study.

One potential area of concern is the composition of the ANES samples. They are high-quality and nationally representative of the US population. They therefore include markedly more participants with less education and internet experience than Project Implicit samples or laboratory studies of university students, and these individuals may have struggled with taking the BIAT or AMP online. We examined the ANES methodology and administration files and located a report of the pilot test of the BIAT in the ANES that noted complaints about the tediousness of the testing procedure and high rates of attrition (Krosnick and Lupia., 2008). We then contacted the ANES staff, however, and learned that they observed no significant problems in the administration of either the BIAT or the AMP in either online or face-to-face interviews (see also DeBell *et al.*, 2010).

A second possible concern is that the BIAT and the AMP were administered as part of a 20-wave panel study. Perhaps the frequency of being interviewed and re-interviewed led to decreases in respondent attentiveness or increases in survey satisficing. Fortunately, the AMP was also administered in the post-election 2008 ANES time series study, which allows us to examine the extent to which respondents perform similarly on the same test administered outside of the panel environment. The results in [Figure 2](#) and [Table I](#) suggest that the same problems we have identified with implicit measures in the panel persist on a different high-quality, nationally representative sample that was arguably less susceptible to respondent fatigue. We can test how the AMP in the panel and the time series compare further by examining the correlation between AMP scores and other key variables that appear in both studies. For example, for one of the most consequential manifestations of political behavior, presidential vote choice, the correlation between AMP score and voting for Barack Obama was -0.14 in both the panel and the time series. The frequency of reinterviewing on the 2008–2009 ANES panel therefore cannot explain the trivial relationship between the BIAT and the AMP.

A final possible concern is that, to fit the IAT into an ANES interview, the shorter version (BIAT), rather than the complete IAT, was administered. Longer measures are more reliable than shorter ones, other things equal. One possibility, then, is that the trivial relationship between the BIAT and the AMP is a function of the lack of reliability of the BIAT. The BIAT is unreliable: the test-retest reliability coefficient is 0.43 (Greenwald and Lai, 2020). However, the complete IAT as a measure of racial prejudice is similarly unreliable, with test-retest reliability for intervals of only 1–2 months of 0.42. The black/white AMP, moreover, has test-reliability coefficients of just 0.35 (Gawronski *et al.*, 2017). While the most robust examination of these concerns would compare the performance of the AMP, the BIAT, and the IAT on a high-quality, nationally representative sample like the one provided by the ANES, we view it as extremely unlikely that such an examination would change our substantive conclusions about the invalidity of implicit measures

for the study of racial prejudice, public opinion, and political behavior. Rather, the sheer magnitude of error in implicit measures of racial prejudice is the most likely explanation for our findings.

It is not obvious how to reconcile such measurement error with the continuing use of measures of implicit prejudice. Indeed, the scale of their usage in racial sensitivity programs as well as academic research on the surface would seem to give a tacit warranty of their predictive validity as indicators of prejudice and discrimination. In fact, what is striking is the lack of predictive power of the BIAT and the AMP. The zero order correlation between measures like the BIAT or the AMP and vote choice is weak (less than 0.2) (Greenwald *et al.*, 2009; Pasek *et al.*, 2009) and consistently trivial after accounting for explicit prejudice (e.g., Payne *et al.*, 2008; Kalmoe and Piston, 2013; Kinder and Ryan, 2017).⁶ Moreover, both critics and proponents agree that the same holds true for the predictive validity of measures of implicit prejudice as barometers of racially aversive or discriminatory responses. Meta-analyses of the relationship between implicit measures and these overt measures of anti-Black hostility report zero order predictive validity coefficients averaging 0.13 (Oswald *et al.*, 2013) or, in a more restricted meta-analysis, 0.26 (Greenwald *et al.*, 2015).

4. Conclusion

Social psychologists have radically updated their understanding of what the BIAT and the AMP measure. Proponents as well as critics now agree that “implicit” processes are not necessarily hidden, unconscious, uncontrolled, automatic, or even implicit (e.g., Fazio and Olson, 2003; De Houwer *et al.*, 2007; Ito *et al.*, 2015; Gawronski *et al.*, 2017; Greenwald and Banaji, 2017; Payne *et al.*, 2017). Our results speak to their specific application in studies of prejudice and politics.

Political scientists who use the BIAT and/or the AMP to study prejudice claim that they circumvent social desirability pressures that lead individuals to consciously or unconsciously hide prejudiced racial attitudes. The ANES is the highest-quality study available to political scientists. It is not only the national representativeness of its sample that sets it apart from other data sources; it also features meticulous development and administration procedures. This is the first study that takes advantage of the ANES to evaluate the validity of the BIAT and the AMP as measures of implicit racial prejudice.

Our results have shown that the relationship between the BIAT and the AMP is substantively indistinguishable from zero. In response, social psychologists might argue that each measures a distinct type of prejudice, which is why the correlation between them is so small. For political scientists, the position that every measure of implicit prejudice measures a different type of prejudice is not very useful for understanding how prejudice affects politics. Moreover, the distinctive virtue of implicit measures of prejudice in political science is supposed to be their power to identify false positives—white people who express positive attitudes toward Black people when they believe that others may learn what they think but who, in reality, dislike and disdain Black people. In fact, our results have shown that the BIAT and the AMP categorize an improbable number of whites as preferring Black Americans relative to white Americans while, simultaneously, classifying as free of anti-Black prejudice substantial numbers of whites who believe that Black people are inferior to white people. These are powerful shortcomings, and all appear to be driven by the unreliability of the measures. To their credit, proponents of implicit measures of prejudice in social psychology have repeatedly assessed the reliability of measures of implicit prejudice. Without exception, test-retest reliability coefficients for implicit measures of racial prejudice hover in the low 0.40s for periods as short as 1–2 months (Gawronski *et al.*, 2017; Greenwald and Lai, 2020).

⁶After accounting for party identification and racial resentment, the standardized coefficient (rescaled from 0–1) obtained from regressing Obama vote on BIAT score is -0.03 ($SE = 0.02$) and on AMP score is -0.08 ($SE = 0.04$).

Our results should not be interpreted as condemning the use of implicit measures per se. Numerous studies have demonstrated the role of implicit political attitudes in predicting the behavior of, for instance, undecided voters (e.g., Lundberg and Payne, 2014; Friese *et al.*, 2016; Ryan, 2017). In a recent innovative study, for example, Ryan and Krupnikov (2021) document how implicit attitudes toward political candidates change in response to emotionally valenced campaign ads. Indeed, social psychologists have demonstrated that implicit measures of political attitudes have test-retest reliability coefficients that run on the order of $r = 0.80$ —twice the size as those of implicit measures of prejudice (Gawronski *et al.*, 2017). The irony is that implicit measures are least satisfactory for the measurement of what many political scientists have presupposed that they are best suited—hidden racial prejudice.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/psrm.2022.56> and replication materials at <https://doi.org/10.7910/DVN/PGPGFH>.

Acknowledgments. We thank Ellen Chapin, Olivier Corneille, Stephen Haber, Andy Hall, Shanto Iyengar, Rachel Lienesch, Hakeem Jefferson, Greg Mitchell, Natalya Rahman, Cole Tanigawa-Lau, Chenoa Yorgason, and anonymous reviewers for helpful comments.

References

- Arkes HR and Tetlock PE (2004) Attributions of implicit prejudice, or “would Jesse Jackson ‘fail’ the implicit association test?”. *Psychological Inquiry* 15, 257–278.
- Bar-Anan Y and Nosek BA (2014) A comparative investigation of seven indirect attitude measures. *Behavior Research Methods* 46, 668–688.
- Chudy J (2021) Racial sympathy and its political consequences. *The Journal of Politics* 83, 122–136.
- Corneille O and Hütter M (2020) Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review* 24, 212–232.
- DeBell M, Krosnick JA and Lupia A (2010) Methodology report and user’s guide for the 2008–2009 ANES Panel Study. https://electionstudies.org/wp-content/uploads/2009/03/ane_s_specialstudy_2008_2009panel_MethodologyRpt.pdf.
- De Houwer J, Beckers T and Moors A (2007) Novel attitudes can be faked on the implicit association test. *Journal of Experimental Social Psychology* 43, 972–978.
- Ditonto TM, Lau RR and Sears DO (2013) AMPing racial attitudes: comparing the power of explicit and implicit racism measures in 2008. *Political Psychology* 34, 487–510.
- Dovidio JF, Kawakami K and Gaertner SL (2002) Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology* 82, 62–68.
- Engelhardt AM (2021) Observational equivalence in explaining attitude change: Have white racial attitudes genuinely changed? *American Journal of Political Science*. doi: 10.1111/ajps.12665.
- Fazio RH and Olson MA (2003) Implicit measures in social cognition research: their meaning and use. *Annual Review of Psychology* 54, 297–327.
- Finn C and Glaser J (2010) Voter affect and the 2008 U.S. presidential election: hope and race mattered. *Analyses of Social Issues and Public Policy* 10, 262–275.
- Friese M, Smith CT, Koeber M and Bluemke M (2016) Implicit measures of attitudes and political voting behavior. *Social and Personality Psychology Compass* 10, 188–201.
- Gawronski B and Bodenhausen GV (2006) Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin* 132, 692–731.
- Gawronski B, Morrison M, Phills CE and Galdi S (2017) Temporal stability of implicit and explicit measures: a longitudinal analysis. *Personality and Social Psychology Bulletin* 43, 300–312.
- Greenwald AG and Banaji MR (2017) The implicit revolution: reconceiving the relation between conscious and unconscious. *American Psychologist* 72, 861–871.
- Greenwald AG and Lai CK (2020) Implicit social cognition. *Annual Review of Psychology* 71, 419–495.
- Greenwald AG, McGhee DE and Schwartz JLK (1998) Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology* 74, 1464–1480.
- Greenwald AG, Nosek BA and Sriram N (2006) Consequential validity of the implicit association test: comment on Blanton and Jaccard (2006). *American Psychologist* 61, 56–61.
- Greenwald AG, Smith CT, Sriram N, B.-A. Y and Nosek BA (2009) Implicit race attitudes predicted vote in the 2008 U.S. presidential election. *Analyses of Social Issues and Public Policy* 9, 241–253.
- Greenwald AG, Poehlman TA, Uhlmann EL and Banaji MR (2009) Understanding and using the implicit association test: iii. meta-analysis of predictive validity. *Journal of Personality and Social Psychology* 97, 17–41.

- Greenwald AG, Banaji MR and Nosek BA** (2015) Statistically small effects of the implicit association test can have societally large effects. *Journal of Personality and Social Psychology* **108**, 553–561.
- Gregg AP, Seibt B and Banaji MR** (2006) Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology* **90**, 1–20.
- Huddy L and Feldman S** (2009) On assessing the political effects of racial prejudice. *Annual Review of Political Science* **12**, 423–47.
- Ito TA, Friedman NP, Bartholow BD, Correll J, Loersch C, Altamirano LJ and Miyake A** (2015) Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology* **108**, 187–218.
- Iyengar S and Westwood SJ** (2015) Fear and loathing across party lines: new evidence on group polarization. *American Journal of Political Science* **59**, 690–707.
- Kalmoe NP and Piston S** (2013) Is implicit prejudice against blacks politically consequential? Evidence from the AMP. *Public Opinion Quarterly* **77**, 305–322.
- Kam CD** (2007) Implicit attitudes, explicit choices: when subliminal priming predicts candidate preference. *Political Behavior* **29**, 343–367.
- Kinder DR and Ryan TJ** (2017) Prejudice and politics re-examined: the political significance of implicit racial bias. *Political Science Research and Methods* **5**, 241–259.
- Krosnick JA and Lupia. A** (2008) Decisions made about implicit attitude measurement in the 2008 American national election studies.
- Lundberg KB and Payne BK** (2014) Decisions among the undecided: implicit attitudes predict future voting behavior of undecided voters. *PloS one* **9**, e105655.
- Malhotra N, Margalit Y and Mo CH** (2013) Economic explanations for opposition to immigration: distinguishing between prevalence and conditional impact. *American Journal of Political Science* **57**, 391–410.
- Messing S, Jabon M and Plaut E** (2016) Bias in the flesh. *Public Opinion Quarterly* **80**, 44–65.
- Mo CH** (2015) The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Political Behavior* **37**, 357–395.
- Oswald FL, Mitchell G, Blanton H, Jaccard J and Tetlock PE** (2013) Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology* **105**, 171–192. /record/2013-20587-001
- Pasek J, Tahk A, Lelkes Y, Krosnick JA, Payne BK, Akhtar O and Tompson T** (2009) Determinants of turnout and candidate choice in the 2008 U.S. presidential election: illuminating the impact of racial prejudice and other considerations. *Public Opinion Quarterly* **73**, 943–994.
- Payne BK, Cheng CM, Govorun O and Stewart BD** (2005) An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology* **89**, 277–293.
- Payne BK, Govorun O and Arbuckle NL** (2008) Automatic attitudes and alcohol: does implicit liking predict drinking?. *Cognition & Emotion* **22**, 238–271.
- Payne BK, Krosnick JA, Pasek J, Lelkes Y, Akhtar O and Tompson T** (2010) Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology* **46**, 367–374.
- Payne BK, Vuletic HA and Lundberg KB** (2017) The bias of crowds: how implicit bias bridges personal and systemic prejudice. *Psychological Inquiry* **28**, 233–248.
- Pérez EO** (2010) Explicit evidence on the import of implicit attitudes: the IAT and immigration policy judgments. *Political Behavior* **32**, 517–545.
- Pérez EO** (2016) *Unspoken Politics: Implicit Attitudes and Political Thinking*. New York: Cambridge University Press.
- Ryan TJ** (2017) How do indifferent voters decide? The political importance of implicit attitudes. *American Journal of Political Science* **61**, 892–907.
- Ryan TJ and Krupnikov Y** (2021) Split feelings: understanding implicit and explicit political persuasion. *American Political Science Review* **115**, 1424–1441.
- Segura GM and Valenzuela AA** (2010) Hope, tropes, and dopes: Hispanic and white racial animus in the 2008 election. *Presidential Studies Quarterly* **40**, 497–514.
- Sriram N and Greenwald AG** (2009) The brief implicit association test. *Experimental Psychology* **56**, 283–294.
- Valentino NA, Neuner FG and Vandenberg LM** (2018) The changing norms of racial political rhetoric and the end of racial priming. *The Journal of Politics* **80**, 757–771.

Appendix A: Explicit prejudice question wording

The question wording for the explicit measures of racial prejudice that were used in the ANES panel study was slightly different in the 2008 ANES time series study. Moreover, the measures in the time series were all administered in the same wave. In the panel study, respondents were asked: “How well does the word ‘lazy’ describe most [whites/Blacks]?” and “How well does the word ‘intelligent at school’ describe most [whites/Blacks]?” on five-point scales (in a random order). In the time series, the questions read: “Now I have some questions about different groups in our society. I’m going to show you a seven-point scale on which the characteristics of the people in a group can be rated. In the first statement a score of 1 means that you think almost all of the people in that group tend to be ‘hard-working.’ A score of 7 means that you think most people in the group are ‘lazy.’ A

score of 4 means that you think that most people in the group are not closer to one end or the other, and of course, you may choose any number in between. . . The next set asks if people in each group tend to be ‘intelligent’ or ‘unintelligent’ . . . Where you rate [WHITES/BLACKS] in general on this scale?” (randomly ordered among a series of other groups).

For the feeling thermometers in the panel, questions asked, “Do you feel warm, cold, or neither warm nor cold toward [whites/Blacks]?” (in a random order) and then asked whether respondents felt “extremely,” “moderately,” or “a little” warm or cold, creating a 7-point composite measure for each group. By contrast, the ANES time series feeling thermometer asked: “I’d like to get your feelings toward some of our political leaders and other people who are in the news these days. I’ll read the name of a person and I’d like you to rate that person using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don’t feel favorable toward the person and that you don’t care too much for that person. You would rate the person at the 50 degree mark if you don’t feel particularly warm or cold toward the person. Still using the thermometer, how would you rate the following groups: [WHITES/BLACKS]” (randomly ordered among a series of other groups).

Appendix B: Additional figure

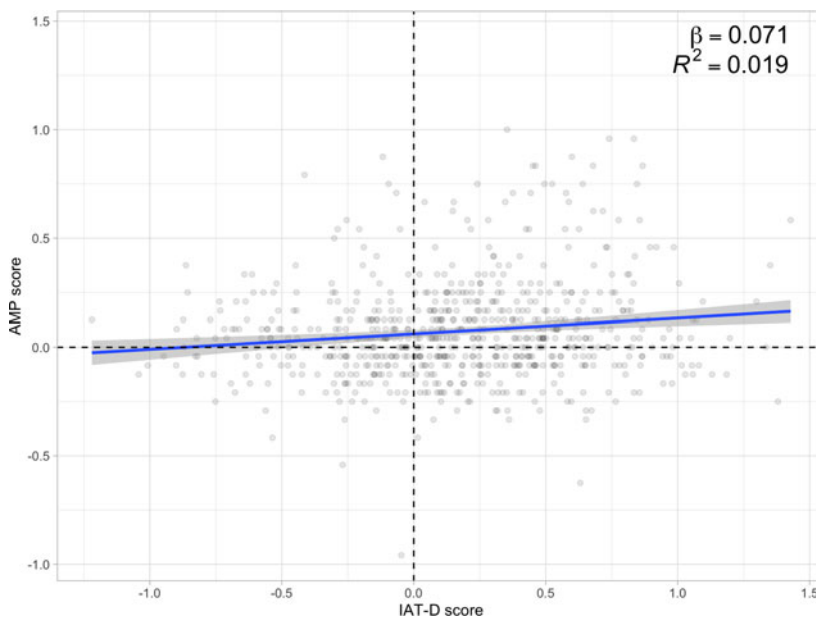


Figure B3. Relationship between BIAT-D scores and AMP scores among Wave 9 Black-white AMP respondents only. Note $N = 679$ white respondents in the 2008–2009 ANES panel study who took the Black-white AMP in Wave 9 *before* the Obama-McCain AMP in Wave 10 (following Kalmoe and Piston, 2013). BIAT-D scores are measured on a -2 to 2 scale, and AMP scores on a -1 to 1 scale (higher values indicating higher anti-Black prejudice). The blue line is an OLS regression line and the shaded area is the 95 percent confidence interval. The results are substantively identical to those observed in Figure 1 in the main text.