The Best Way to Measure Explicit Racial Attitudes Is to Ask About Them

Jordan R. Axt[1]

Abstract

Direct assessments of explicit racial attitudes, such as reporting an overt preference for White versus Black people, may raise social desirability concerns and reduce measurement quality. As a result, researchers have developed more indirect self-report measures of explicit racial attitudes. While such measures dampen social desirability concerns, they may weaken measurement quality by assessing construct-irrelevant attitudes, thereby lowering correspondence between measure and construct. To investigate whether direct or indirect self-report measures better assess explicit racial attitudes, participants (*N* > 800,000) completed an implicit racial attitude measure and a subset of over 400 items that varied in the degree to which they were indirect or direct assessments of self-reported racial attitudes. More direct assessments of racial preferences were better predictors of implicit racial attitudes and maximized differences between Black and White participants. These results suggest that the best method to measure individuals' explicit racial attitudes is to ask about them directly.

Keywords

attitudes, social cognition, dual process models, measurement, prejudice/stereotyping, self-presentation

Few dissertations had a larger impact on psychological research than one in 1932 by a former seminary student. The project proposed that attitudes could be assessed by having people respond to related items, assign numeric values to response options, and average responses. Rensis Likert's approach proved effective, efficient, and reliable, and psychological measurement has not been the same since.

Likert (1932) investigated attitudes regarding issues like economic conflict and race relations. His questions concerning White–Black attitudes, called the Negro scale, were particularly straightforward. Participants answered items like "Would you shake hands with a Negro?" or indicated the extent to which they agreed with the statement, "Negro homes should be segregated from those of White people."

While those items may have been useful in 1932, researchers today may hesitate to use similar items, citing shifts in public attitudes and changes in norms about expressing prejudice. Such items can be considered highly direct measures of self-reported racial attitudes (here, meaning evaluations of Black vs. White people). In this sense, "directness" does not refer to measurement procedure, such as differentiating between measures capturing more controlled responses, like self-report scales, versus more automatic responses, like the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). Rather, "directness" in this context refers to the degree of correspondence between the measure's content and the target construct, meaning two self-report items can differ on their degree of directness. Self-report items containing

self-assessment of the target construct are more direct (De Houwer & Moors, 2010), whereas self-report items containing self-assessment of objects only related to the target construct are more indirect. However, indirect self-report items are still considered measures of explicit rather than implicit attitudes, as responses are more influenced by conscious goals (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009) and more intentional (Moors & De Houwer, 2006) compared to responses on implicit measures like the IAT.

For racial attitudes, highly direct self-report items may introduce social desirability concerns (Paulhus, 1984), prompting participants to not report their real attitudes but socially acceptable responses, thereby weakening measurement of the construct of explicit racial attitudes. One approach to mitigating desirability concerns has been through creating more indirect self-report measures, where participants express negative racial attitudes but can justify responses by referring to external factors that do not necessitate prejudice. For example, participants harboring anti-Black attitudes may be willing to disagree with an indirect measure of self-reported racial attitudes like "Sometimes Black job seekers should be given

---

[1] Department of Psychology, University of Virginia, Charlottesville, VA, USA

Corresponding Author:
Jordan R. Axt, Department of Psychology, University of Virginia, Box 400400, Charlottesville, VA 22904, USA.
Email: jra3ee@virginia.edu

special consideration in hiring" (Katz & Hass, 1988) because such an attitude can be plausibly justified by a belief that no race should be given preferential treatment. However, those same participants may be less willing to agree with a more direct measure of self-reported racial attitudes, such as whether Black people are lazy (Peffley, Hurwitz, & Sniderman, 1997), because such an opinion can only indicate negativity toward Black people.

In the last 35 years, a number of indirect self-report measures of racial attitudes have been developed, such as Symbolic Racism (Sears, 1988), Modern Racism (McConahay, 1986), and New Racism (Jacobson, 1985). While such scales are direct measures of the race-related policies or perceptions assessed by their items, they are commonly used as measures of explicit racial attitudes (e.g., Blincoe & Harris, 2009; Derous, Nguyen, & Ryan, 2009; Goff, Steele, & Davies, 2008; Groom, Bailenson, & Nass, 2009; Guinote, Willis, & Martellotta, 2010; Gulker & Monteith, 2013; Huntsinger & Smith, 2009; Inzlicht, Gutsell, & Legault, 2012; Legault, Gutsell, & Inzlicht, 2011; Lybarger & Monteith, 2011; Olson & Fazio, 2006; Rudman & Lee, 2002; Stanley, Sokol-Hessner, Banaji, & Phelps, 2011; Trawalter, Hoffman, & Waytz, 2012).

Researchers may assume that indirect self-report measures better assess explicit racial attitudes due to lower social desirability concerns. However, while indirect self-report measures can lessen desirability concerns, they also reduce correspondence or similarity between measure and construct (Ajzen & Fishbein, 1977). By definition, indirect measures of self-reported racial attitudes introduce construct-irrelevant information; for instance, opinions toward affirmative action involve construct-relevant components, like racial preferences, but also construct-irrelevant components, like thoughts on educational policy.

Lower correspondence could result in indirect self-report items being less valid measures of explicit racial attitudes compared to more direct items. Indeed, meta-analyses have found that increased conceptual correspondence between measures was associated with larger correlations between implicit and explicit attitudes (Hofmann et al., 2005) and between implicit attitudes and behavior (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). However, in cases like the Hofmann et al. meta-analysis, such findings are difficult to interpret when estimates collapse across attitude domains, as it's possible that studies in areas with higher implicit–explicit correlations, such as consumer attitudes (e.g., Maison, Greenwald, & Bruin, 2001), may have also been more likely to use direct self-report measures.

Indirect self-reports of racial attitudes are then low on correspondence, which may weaken construct measurement, but are simultaneously low on social desirability, which may strengthen construct measurement. The current analysis investigates this trade-off by examining whether more direct or indirect self-report items are better measures of explicit racial attitudes. Online participants completed a subset of over 400 self-report items that varied in the extent to which they were direct measures of racial attitudes. More direct versus more indirect measures of self-reported racial attitudes were compared on (1) correlation with a Black–White evaluative IAT (Greenwald et al., 1998) and (2) mean difference in responses between Black and White participants.

Better measures of a construct will elicit stronger relations with related measures than worse measures. Height and weight are distinct but correlated constructs, and measures that best reduce random error will maximize the correlation between the two and get closest to their true correlation (see Nosek, Bar-Anan, Sriram, Axt, & Greenwald, 2014, for parallel reasoning). Implicit and explicit racial attitudes have been repeatedly shown to correlate ($r = .31$, $N > 180,000$ in Schmidt & Axt, 2016), meaning the best measures of explicit racial attitudes should maximize the true correlation between the two constructs. Just as in height and weight, this logic applies whether one conceives of explicit and implicit evaluations as measuring the same construct or separate but related constructs, the latter position being well established (Nosek & Smyth, 2007). Similarly, there are group differences in explicit racial attitudes ($d = 1.29$ between Black and White participants in Schmidt & Axt, 2016). These differences will be weakened by measurement error, except in unlikely cases where error is confounded with racial identity (see Greenwald, Nosek, & Banaji, 2003). Minimizing measurement error will maximize the magnitude of these differences and bring them closer to their true value, meaning better measures of self-reported racial attitudes should maximize differences between White and Black respondents.

This investigation represents one analysis of a larger data collection concerning relations among racial attitude measures. The sample contains over 800,000 participants who completed a measure of implicit racial attitudes and a subset of over 400 self-report items. Analyses focus on the role of item directness in assessing the construct of explicit racial attitudes. While self-report items rated by coders to be more direct assessments of racial attitudes were also rated as more likely to create socially desirable responding, these more direct self-report items still produced larger correlations with the IAT and greater mean differences between Black and White respondents. Despite greater potential for socially desirable responding, more direct self-report items better assess the construct of explicit racial attitudes.

Method

Participants

Participants were volunteers who selected the Race IAT at Project Implicit (implicit.harvard.edu). Data collection occurred between October 23, 2014, and September 27, 2016. End date was selected arbitrarily when the study was updated. During that time, 1,396,234 participants provided data ($M_{\text{age}} = 27.3$, $SD_{\text{age}} = 12.2$; 60.1% female, 68.5% White, 9.7% Black; 82.8% U.S. residents).

Materials

Explicit racial attitudes. Thirty-four self-report scales measuring racial or political attitudes were used, 31 from previous research and 3 developed for data collection (see Table 1

Table 1. Reference, Number of Items, Number of Participants, and Mean Dimension Ratings for Each Scale.

Scale (Abbreviation)	Source	Items	Responses	Feel	Personal	Directness	Elaboration	Desirability
ANES: Race[a] (ANES)	Payne et al. (2010)	6	46,746	1.07	1.0	1.76	1.32	1.93
Attitudes Toward Blacks (ATB)	Brigham (1993)	20	45,605	0.65	1.0	1.45	1.48	1.75
Attitudes Toward Whites (ATW)	Brigham (1993)	20	44,714	0.44	0.89	1.41	1.29	1.76
Bayesian Racism (BR)	Uhlmann, Brescoll, and Machery (2010)	6	47,476	0.03	1.0	1.03	1.57	2.15
Classical Conservatism (CC)	McClosky (1958)	9	48,134	0.19	1.0	0.04	1.07	0.44
Explanations for Poverty	Kluegel and Smith (1986)	12	44,752	0.17	1.0	0.02	1.55	1.49
Individual Explanations (EFP-I)	—	5	47,444	0.08	1.0	0.05	1.52	1.42
Structural Explanations (EFP-S)	—	5	47,468	0.23	1.0	0	1.75	1.68
General Intergroup Anxiety (GIA)	Stephan and Stephan (1985)	12	48,782	1.99	1.0	1.89	1.48	1.42
GSS: Opportunity (GSS-Opp)	Davis and Smith (1990)	9	47,819	0.18	1.0	1.12	1.49	1.91
GSS: Race (GSS-Race)	Davis and Smith (1990)	22	43,317	0.50	1.0	1.58	1.28	2.02
Intergroup Anxiety (IA)	Britt, Bonieci, Vescio, Biernat, and Brown (1996)	11	46,258	1.48	1.0	1.64	1.24	1.66
Motivation to Control Prejudice	Plant and Devine (1998)	10	47,161	0.63	1.0	0.52	1.72	1.57
Internal Motivation (IMS)	—	5	47,655	0.37	1.0	0.55	1.67	1.58
External Motivation (EMS)	—	5	47,375	0.90	1.0	0.48	1.77	1.55
Motivation to Control Prejudiced Responses	Dunton and Fazio (1997)	17	46,517	0.78	0.94	0.45	1.52	1.23
Concerns With Acting Prejudiced (MCPR-C)	—	9	46,712	0.75	0.89	0.64	1.63	1.38
Restraint To Avoid Dispute (MCPR-R)	—	6	47,167	0.81	1.0	0.31	1.31	1.06
Moral Conservatism (MC)	Wald, Owen, and Hill (1988)	10	48,920	0.09	1.0	0.16	1.91	2.07
Modern Racism (MR)	McConahay (1986)	7	48,817	0.14	1.0	1.47	1.26	2.06
New Racism (NR)	Jacobson (1985)	7	45,712	0.39	1.0	1.39	1.29	1.93
Pro-Black Anti-Black Attitudes	Katz and Hass (1988)	20	45,805	0.02	1.0	1.35	1.32	1.97
Anti-Black Attitudes (ABA)	—	10	46,318	0.01	1.0	1.46	1.07	1.93
Pro-Black Attitudes (PBA)	—	10	46,456	0.03	1.0	1.25	1.56	2.02
Perceived Group Conflict[b] (PGC)	Sidanius, Van Laar, Levin, and Sinclair (2004)	5	46,780	0.30	0.80	0.77	1.95	1.98
Political Tolerance (PT)	Duckitt and Farre (1994)	6	45,876	0	1.0	0	1.82	2.46
Prejudice Index (PI)	Bobo and Kluegel (1993)	10	46,708	0.07	1.0	1.83	1.28	2.03
Racial Ambivalence: Black	Katz and Hass (1988)	20	44,821	0.03	1.0	1.28	1.22	1.95
Anti-Black Attitudes (RA-ABA)	—	10	45,405	0.02	1.0	1.37	1.08	1.90
Pro-Black Attitudes (RA-PBA)	—	10	45,608	0.05	1.0	1.20	1.36	1.99
Racial Ambivalence: Work	Katz and Hass (1988)	21	46,475	0.09	1.0	0.02	1.45	1.09
Protestant Ethic (RA-PE)	—	11	46,800	0.16	1.0	0.03	1.23	.77
Humanitarianism–Egalitarianism (RA-HE)	—	10	46,837	0.01	1.0	0	1.68	1.43
Racial Arguments (RaceArg)	Saucier and Miller (2003)	13	41,030	0.08	1.0	0.97	1.04	1.56
Racial Attitudes (RA)	Sidanius, Pratto, Martin, and Stallworth (1991)	14	47,052	1.29	1.0	1.47	1.55	2.33
Racial Resentment (RR)	Kinder and Sanders (1996)	6	48,567	0	1.0	1.40	1.31	2.04
Racial Stereotypes Measure (RSM)	Peffley, Hurwitz, and Sniderman (1997)	6	46,967	0.31	1.0	1.92	1.35	2.03
Right-Wing Authoritarianism (RWA)	Altemeyer (1981)	20	42,833	0.02	1.0	0	1.33	1.71
Subtle and Blatant Prejudice	Pettigrew and Meertens (1995)	20	44,065	0.35	1.0	1.50	1.20	2.07
Subtle Prejudice (SP)	—	10	44,916	0.40	1.0	1.46	1.13	1.82
Blatant Prejudice (BP)	—	10	45,851	0.30	1.0	1.53	1.28	2.31
Social Dominance Orientation (SDO-6)	Sidanius and Pratto (2001)	16	47,192	0.05	1.0	0.35	1.57	2.25
Symbolic Racism 2000 (SR2000)	Henry and Sears (2002)	8	46,282	0.15	1.0	1.33	1.27	1.85
Trust in People[c] (TIP)	Robinson, Shaver, and Wrightsman (1999)	3	46,896	0	1.0	0	1.81	1.31
Universal Orientation (UO)	Phillips and Ziller (1997)	20	45,379	0.17	1.0	0	1.32	0.71

(continued)

Table 1. (continued)

Scale (Abbreviation)	Source	Items	Responses	Feel	Personal	Directness	Elaboration	Desirability
Cultural Attitudes Toward Black People[c] (CAB)	—	6	48,623	1.24	0	1.57	1.79	2.33
Cultural Attitudes Toward White People[b] (CAW)	—	6	46,850	1.64	0	1.57	1.36	1.68
Others' Preferences[a] (OP)	—	6	47,788	0.43	0	1.58	1.75	2.33

Note. ANES = American National Election Survey; GSS = General Social Survey. Feel = ratings of beliefs versus feelings; personal = ratings of cultural versus personal attitudes; directness = ratings of indirect versus direct measurement; elaboration = ratings of low versus high elaboration; desirability = ratings of low versus high desirability. Data are presented for both the full scale and any subscales. For the Explanations for Poverty and the Motivations to Control Prejudiced Responses Scales, the full scale contains items that are not included in any subscales. Scales sharing a superscript were administered together.

for a list of scales, their source, number of items, and responses). Shorter scales were combined into one set for administration, resulting in 31 item sets ($M = 13.2$ items, $SD = 5.30$). Participants received two sets selected at random, or one set and another questionnaire about life experiences not analyzed here.

Several changes were made to scales. First, items referring to specific places were changed to "my country." Second, "Blacks" and "Whites" were changed to "Black people" and "White people"; one scale (Pettigrew & Meertens, 1995) about West Indian versus British attitudes was updated to be about Black and White people. Third, a neutral "neither agree or disagree" response was added to agreement items. Finally, word descriptions were added to response options that were only numeric, a change associated with greater reliability (Alwin & Krosnick, 1991). Wording changes are available in the Online Supplement at https://osf.io/e9shx.

Participants completed 3 items directly before or after the first item set. First, participants completed an explicit racial preference item asking "Which statement best describes you?" ($-3 = I$ *strongly prefer African Americans to European Americans*, $0 = I$ *like European Americans and African Americans equally*, $+3 = I$ *strongly prefer European Americans to African Americans*). Next, participants completed two thermometer items assessing warmth toward African Americans and European Americans separately ($1 = extremely\ cold$, $11 = extremely\ warm$).

Race IAT. Participants completed a seven-block race IAT (Greenwald et al., 1998) measuring strength of association between the concepts "Good" and "Bad" and the categories "African Americans" and "European Americans." Six images of faces (three male, three female) represented each racial category (available in Online Supplement along with words representing each concept). Procedure followed the design from Nosek, Greenwald, and Banaji (2007) and was scored by the D algorithm (Greenwald et al., 2003). More positive scores reflected more positive associations with European Americans versus African Americans. For analyses involving the IAT, participants having more than 10% of trials faster than 300 milliseconds were excluded (1.6%; Nosek et al., 2007).

Demographics. Demographics were updated during data collection. Participants before March 2, 2015, completed a 15-item survey, and participants afterward completed a 28-item survey. Only responses for age, gender, race, ethnicity, and country of citizenship or residence were analyzed. All demographic variables are available in the online data set.

Procedure

Participants completed the two self-report attitude measures, the demographics survey, and the race IAT in a randomized order.

Coding Explicit Attitude Items

Twelve research assistants rated 400 of the 407 self-report items[1] on six dimensions previously identified as potential moderators of the relationship between implicit and explicit attitudes (e.g., Hofmann, Gschwender, Nosek, & Schmitt, 2005; Nosek, 2005, 2007) or could be used in new tests for implicit–explicit moderators (e.g., whether an item assesses feelings vs. beliefs). This investigation focuses on the role of item directness, but all variables are described to provide a full description of the data set. Items were rated one at a time on each dimension in the following order.

Racial content. Coders rated whether items were about (1) only Black people, (2) only White people, (3) Black and White people, (4) people from other races (specified), (5) race but racial groups not specified, or (6) not about race.

Feelings or beliefs. Coders rated whether items were about (0) beliefs, (1) both feelings and beliefs, or (2) feelings.

Personal or cultural views. Coders rated whether items assessed (0) cultural views or (1) personal views.

Directness. Coders rated whether items measured racial attitudes or preferences directly or indirectly: (0) racial attitudes or preferences not assessed at all, (1) indirect assessment of racial attitudes or preferences, and (2) direct assessment of racial attitudes or preferences.
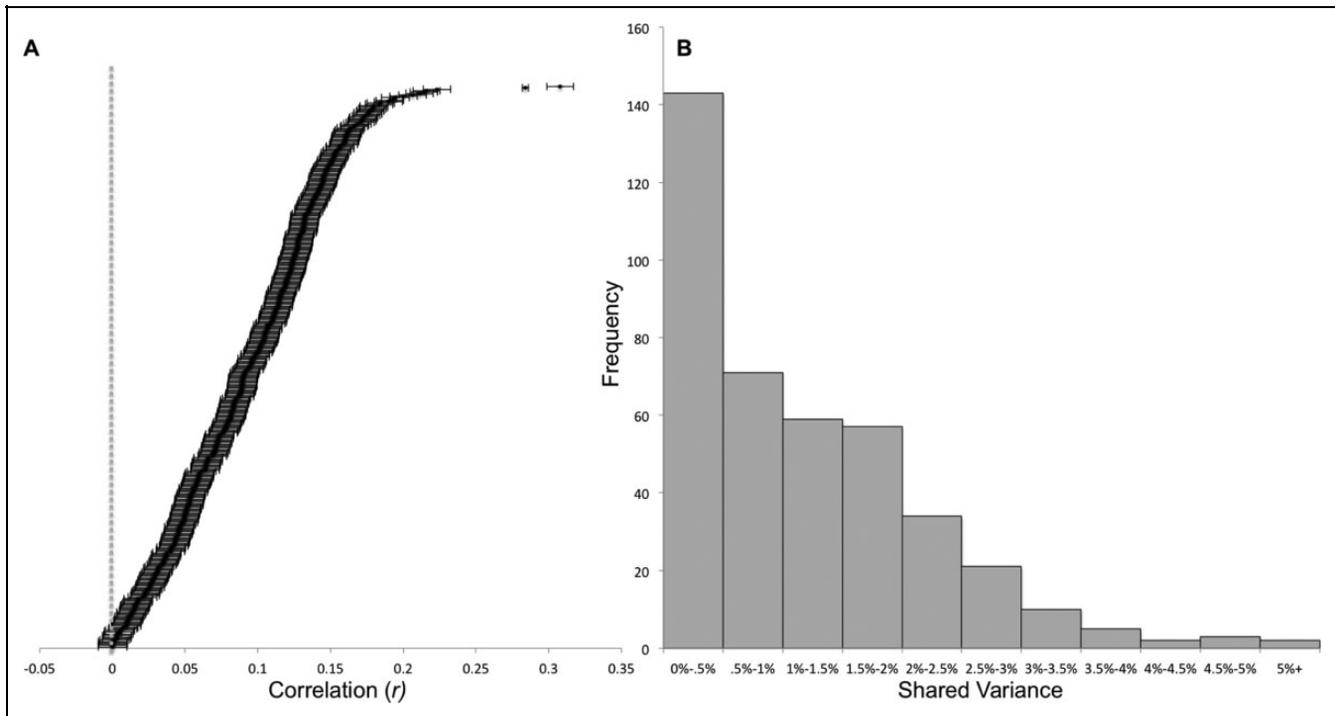
Figure 1. Panel A shows a rank order of all 407 items in strength of correlation with the Implicit Association Test (IAT), with dotted line at 0. Error bands represent 95% confidence intervals. Panel B is a histogram for items' shared variance with the IAT.

Instructions clarified that direct self-reports "ask the respondent to report their feelings, attitudes, or preferences about race," whereas indirect self-reports "may be influenced by racial attitudes, but the response itself is supposedly about something else." Coders were instructed that one way to differentiate direct from indirect self-reports was to ask, "Can someone respond to this question in any way and reasonably argue that they did not use their racial attitudes to influence their response?"

Elaboration. Coders rated the extent to which someone had thought about the opinion assessed: (0) opinion someone has never or very rarely thought about, (1) opinion someone has thought about a little bit, (2) opinion someone has thought about a fair amount, or (3) opinion someone has thought about a great deal.

Social desirability. Coders rated whether a person would be concerned about the social consequences of publicly admitting any of the possible responses: (0) response options produce no social concerns, (1) response options produce a little bit of social concern, (2) response options produce some social concerns, or (3) response options produce a lot of social concerns. Coders did not rate their own or the most common response, but rather whether *any* of the possible responses would bring social consequences.

An intraclass correlation coefficient using a two-way random model indicated good interrater reliability (Hallgren, 2012): racial content = .996, feelings or beliefs = .951, personal or cultural views = .771, directness = .969, elaboration = .747, and social desirability = .869.

Results

Descriptive Statistics

IAT D scores ($N = 930,639$; $M = .28$, $SD = .44$, $d = .64$) and explicit preferences ($N = 1,031,207$; $M = .21$, $SD = .96$, $d = .22$) were both positive. Participants reported slightly higher warmth toward European ($N = 1,041,605$; $M = 7.64$, $SD = 2.03$) than African Americans ($N = 1,042,307$; $M = 7.61$, $SD = 2.07$). White participants reported higher warmth toward European ($N = 541,051$; $M = 7.77$, $SD = 1.97$) than African Americans ($N = 541,258$; $M = 7.47$, $SD = 2.00$), and Black participants reported higher warmth toward African ($N = 74,534$; $M = 9.08$, $SD = 2.01$) than European ($N = 74,497$; $M = 7.78$, $SD = 2.25$) Americans.

There was an average of 39,254 responses with usable IAT scores for each self-report item (minimum = 35,877). All self-report measures were recoded such that positive correlations with the IAT indicated consistency in racial attitudes. Some items did not indicate racial attitudes, leaving ambiguous the expected correlation. Rather than apply theoretical predictions, all self-report measures were coded so that correlation with the IAT was positive (see Figure 1 for a rank ordering of items by correlation with the IAT). The median correlation was $r = .094$.

Figure 1 shows two outlier items. The item most highly correlated with the IAT ($N = 37,618$; $r = .308$, 95% CI [.299, .317]) was from the General Social Survey (GSS) section about
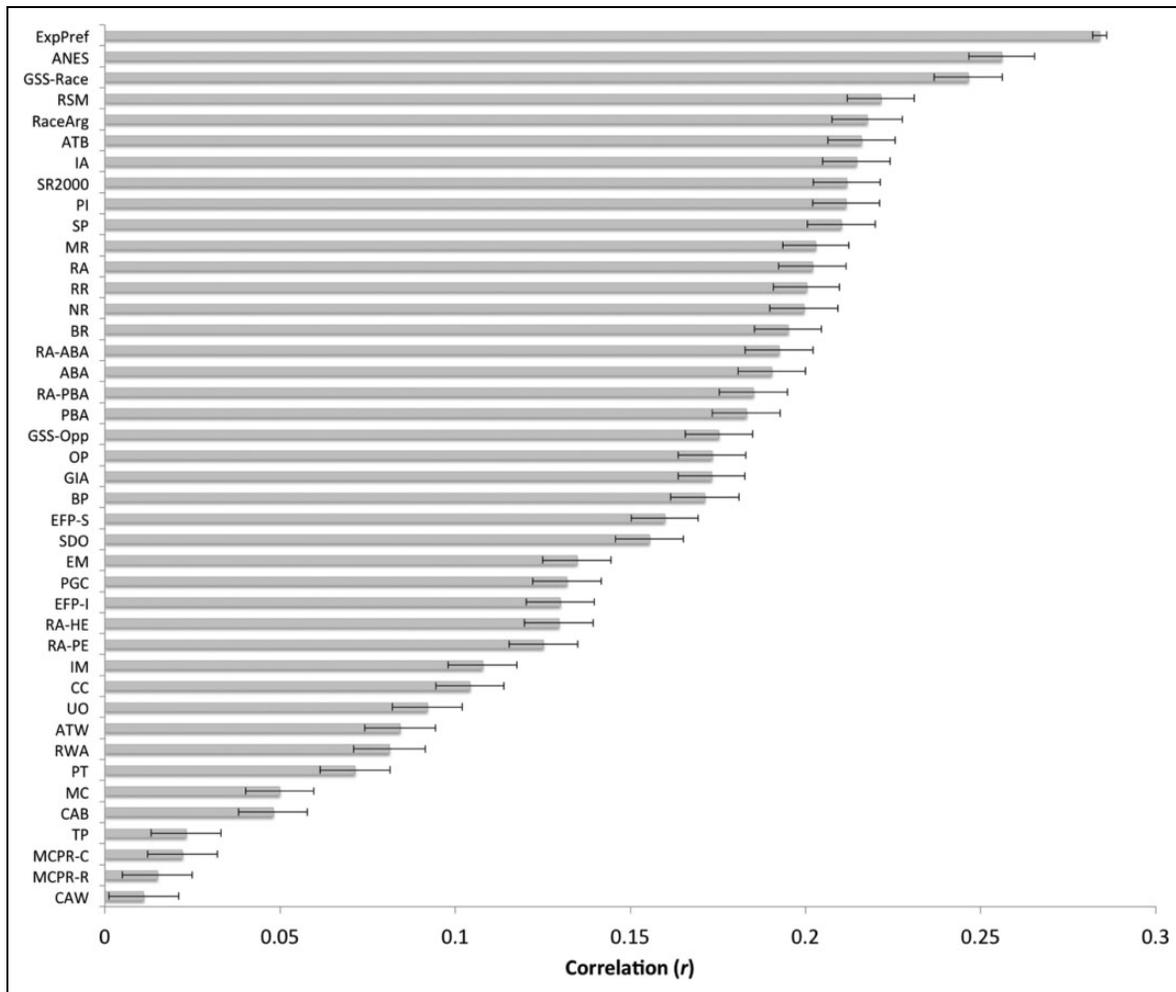
Figure 2. Rank order of each scale or subscale for correlation with the Implicit Association Test (see Table 1 for list of abbreviations). Exp Pref = Single-item explicit racial preference measure. Error bars represent 95% confidence intervals.

race (Race; Davis & Smith, 1991). The item, echoing Likert's own question about segregation, asked "If you could find the housing that you would want and like, would you rather live in a neighborhood that is all Black; mostly Black; half Black; half White; mostly White; or all White?" The other outlier was the explicit preference item concerning how much participants liked European versus African Americans ($N = 826,102$; $r = .284$, 95% CI [.282, .286]).

Aggregate scores for each scale or subscale were also computed (see Figure 2).[2] Across 41 scales or subscales, the median correlation was $r = .173$. The scales with the highest IAT correlation were the race-related items from the GSS ($N = 35,427$; $r = .246$, 95% CI [.237, .256]) and the American National Election Studies ($N = 38,113$; $r = .256$, 95% CI [.247, .265]), though neither was more strongly correlated with the IAT than the single explicit preference item.

Item Directness

The first analysis investigated the relationship between the proportion of shared variance (r^2 between each item and IAT D score) with average ratings of directness. Across items, there was a reliable, positive correlation between directness and proportion of IAT D score variance explained ($r = .474$, $p < .001$, 95% CI [.395, .546]).

However, it's possible that the relationship found among all items could be created through poor predictive performance of items not generally rated as indirect or direct measures of self-reported racial attitudes or of items not concerning Black or White people, and directness may not predict IAT variance when excluding such items. To explore the robustness of this effect, the correlation between directness and IAT variance was also examined among (1) items with directness rating greater than 0, (2) items with directness rating of at least 1, and (3) items dealing only with Black and/or White people.

The reliable relationship between directness and IAT variance weakened but held among items with a directness rating greater than 0 ($N = 307$), meaning at least one coder thought the item directly or indirectly measured racial preferences, $r = .331$, $p < .001$, 95% CI [.228, .427]. The relationship also held among items with directness rating of at least 1 ($N = 245$; $r = .182$, $p = .004$, 95% C.I. [.058, .300]), and among

items that at least 83% of coders rated as concerning Black and/or White people ($N = 230$; $r = .204$, $p = .002$, 95% CI [.077, .324]). Interaction analyses through linear regression illustrated that these subsets of items did not reliably differ in the relationship between directness and IAT variance (all Bs < .01, all ps > .969, see Online Supplement for full reporting).

As expected, there was a positive correlation between directness and social desirability concerns, across all items ($r = .454$, $p < .001$, 95% CI [.373, .528]) and items concerning Black and/or White people ($r = .228$, $p < .001$, 95% CI [.205, .383]). In a series of linear regressions, directness remained a reliable predictor of IAT variance after controlling for social desirability, among all items (directness $B = .43$, $p < .001$; desirability $B = .10$, $p = .037$), items with directness greater than 0 (directness $B = .30$, $p < .001$; desirability $B = .12$, $p = .035$), items with directness of at least 1 (directness $B = .17$, $p = .009$; desirability $B = .10$, $p = .120$), and items concerning Black and/or White people (directness $B = .18$, $p = .008$; desirability $B = .11$, $p = .091$). These results suggest that, beyond raising social desirability concerns, other aspects of more direct self-report items are responsible for explaining greater IAT variance (e.g., more direct items contain more information related to the construct of racial attitudes).

Differences Between Black and White Participants

The second analysis investigated the relationship between directness and differences in average responses from Black versus White participants (Cohen's d comparing White and Black respondents). Three hundred and sixty-six items (89.9%) showed reliable differences between White and Black participants, though many were small effects (median $d = .195$). Taking the absolute value of the Cohen's d for each item, more direct items were associated with greater differences between Black and White participants, $r = .237$, $p < .001$, 95% CI [.143, .327].

This relationship was weaker but reliable for items with a directness rating greater than 0, $r = .150$, $p = .008$, 95% CI [.039, .257]. However, the association between directness and differences between Black and White participants was positive but not reliable among items with directness rating of at least 1 ($r = .079$, $p = .215$, 95% CI [−.046, .202]) or items concerning Black and/or White people ($r = .111$, $p = .092$, 95% CI [−.018, .236]). Interaction analyses illustrated that these subsets of items did not reliably differ in the relationship between directness and Black–White mean differences (all Bs < .38, all ps > .172, see Online Supplement).

Directness remained a reliable predictor of differences between Black and White participants after controlling for social desirability, among all items (directness $B = .20$, $p < .001$; desirability $B = .08$, $p = .136$), and items with directness greater than 0 (directness $B = .13$, $p = .025$; desirability $B = .06$, $p = .342$), but not among items with directness of at least 1 (directness $B = .07$, $p = .281$; desirability $B = .07$, $p = .306$), or items concerning Black and/or White people (directness $B = .09$, $p = .171$; desirability $B = .08$, $p = .231$).

Participant Dropout

Given the volunteer sample and unequal number of items across sets, differential dropout was likely. Indeed, there was a negative correlation ($r = −.54$) between number of participant responses and number of items in a set, suggesting longer scales created more dropout, potentially biasing results if certain types of participants (e.g., those high or low in implicit or explicit racial preferences) were more or less likely to complete the study.

While this possibility cannot be ruled out, there were only minor differences between participants who completed one versus both of the implicit or explicit attitude measures. Among participants completing the IAT, there was a small difference in implicit attitudes between those who reported explicit attitudes ($M = .28$, $SD = .44$) and those who did not ($M = .29$, $SD = .43$; comparison $d = .017$). Conversely, among participants completing the explicit preference item, there were no reliable differences in explicit attitudes between those who completed the IAT ($M = .211$, $SD = .94$) and those who did not ($M = .213$, $SD = 1.02$, comparison $d = .002$). Similar results emerged across the scales and subscales in Figure 2. Among those completing any scale or subscale, there were small differences between participants who did versus did not complete the IAT (average $d = .10$, median $d = .09$) or did versus did not complete the explicit preference item (average $d = .08$, median $d = .08$).

Moreover, there was little dropout once participants completed any self-report item. The median percentage of participants who completed any scale after providing at least one response was 97.5% (average = 96.4%, minimum = 87.9%). These results indicate that little dropout happened once participants completed a single item and suggest there were no substantive attitudinal differences between participants who did or did not complete all study measures.

General Discussion

Across 400 self-report items, more direct measures of self-reported racial attitudes were associated with higher correlations with a race IAT and larger response differences between White and Black participants, despite such items also being associated with greater social desirability concerns. These results align with earlier meta-analytic estimates of correlations between self-report items and the IAT (Hofman, Gawronski, et al., 2005), where greater conceptual correspondence between measures was associated with higher implicit–explicit correlations; more direct self-report items like feeling thermometers ($r = .236$) had stronger correlations than less direct items like scales ($r = .177$). However, this analysis collapsed across seven attitudinal domains, meaning it's possible that researchers in areas with naturally stronger implicit–explicit correlations were also more likely to use direct than indirect self-report measures. By keeping IAT and domain constant, the current work complements these previous findings and provides stronger evidence for the benefits of using direct self-report items to assess explicit attitudes.

These results also align with work showing relatively high percentages of participants willing to self-report beliefs or motives that may be considered socially unacceptable. In research on motivations to express prejudice (Forscher, Cox, Graetz, & Devine, 2015), 21.5% of a sample of undergraduates were neutral or agreed with the statement "Avoiding interactions with Blacks is important to my self-concept". Likewise, in work on blatant dehumanization (Kteily, Bruneau, Waytz, & Cotterill, 2015), more than half of an online sample reported that Arabs and Muslims were less human than Americans or Europeans. In the current data, while a majority indicated no explicit racial preference, 38% still reported a racial preference (28% favored White people, 10% favored Black people). These findings highlight that at least some people are quite willing to express preferences, beliefs, and motivations that privilege some groups over others, and that perhaps concerns over social desirability in self-report measures have been overstated.

However, these data do not suggest that social desirability should be ignored as a threat when assessing self-reported attitudes. Rather, this work extends prior research on the impact of social desirability across attitude domains. Nosek (2005) compared implicit and explicit correlations across 57 attitude objects, adapting the same implicit (IAT) and explicit measure (a feeling thermometer difference score) for all topics. Domains rated as lower in self-presentation concerns had higher implicit–explicit correlations; low self-presentation domains like Coke versus Pepsi ($r = .54$) produced stronger correlations than high self-presentation domains like race (White vs. Black people: $r = .33$). As IAT performance is largely outside of conscious control (Nosek, Banaji, & Greenwald, 2002), such results suggest that participants reporting attitudes high in desirability concerns may have misreported their actual explicit attitudes to provide more socially acceptable responses.

Integrating the current results with those in Nosek (2005) presents an intriguing paradox. *Across* attitude domains, direct self-report measures will have weaker construct measurement in topics high versus low in social desirability. However, *within* an attitude domain, self-report measures will have stronger construct measurement when items are high versus low in directness. The social desirability concerns occasionally created by direct measures in domains like race appear to suppress construct measurement for self-reported attitudes but do so less than the construct-irrelevant information introduced by using more indirect self-report measures.

The Value of Implicit Measures

While direct self-reports of racial attitudes were best correlated with implicit racial attitudes, this does not mean the two measures are redundant. The implicit–explicit correlation found here ($r = .28$) was modest and similar to estimates from other online samples ($r = .33$ in Schmidt & Nosek, 2010), and previous work indicates that implicit and explicit attitudes are best understood as separate but related constructs; for instance, a two-factor correlated model was a superior fit than

a single-factor model for 56 of 57 attitude domains (Nosek & Smyth, 2007).

Moreover, theoretical models argue for distinct sources of information forming implicit and explicit evaluations (e.g., Friese, Hofmann, & Schmitt, 2009; Gawronski & Bodenhausen, 2006), and empirical studies suggest implicit and explicit attitudes are differentially correlated with certain behaviors (Dovidio, Kawakami, & Gaertner, 2002; Hofmann, Gschwendner, Castelli, & Schmitt, 2008) or change at differing rates (Cao & Banaji, 2016; Ranganath & Nosek, 2008). The present results do not challenge whether implicit and explicit attitudes are distinct constructs; they merely suggest that for racial attitudes at least, the best means of measuring explicit attitudes and thereby maximizing the relationship between the two constructs is through the use of direct over indirect self-report items.

Limitations of the Study

Although the sample was large, it was not representative of any definable population. It is possible that representative samples would not show the relationships found here, though I cannot identify a plausible reason to expect this lack of generalizability. Similarly, data came exclusively from online participants, and results may have differed with in-lab samples; participants may have felt greater social desirability concerns or expressed different levels of implicit or explicit racial attitudes in a lab context (e.g., Joy-Gaba & Nosek, 2010, study 3). Investigating whether more direct self-report items continue to exhibit the same measurement advantage when responses are made in the presence of others will be very informative.

It is possible that results were distorted by differential dropout, though dropout was low once participants completed any item and there were only small attitudinal differences between participants who did versus did not drop out. Finally, this work only dealt with one prominent racial attitude, evaluations of Black versus White people. It will be useful to extend this research to other racial attitudes, such as evaluations of other racial groups, or to other racial issues, like multiculturalism, in addition to identifying whether the same relationship between item directness and assessing the construct of explicit attitudes emerges in other domains.

Conclusion

These data contain more than 800,000 responses to racial attitude measures, including more than 35,000 responses to over 40 commonly used scales. Moreover, since participants were assigned multiple scales, there are over 700 responses between any two scales in Table 1. This analysis was focused on the role of item directness in measuring the construct of explicit racial attitudes. However, this is a rich data set, and interested researchers should use these data for additional investigations. All data, materials, analysis syntax, and codebooks are available at https://osf.io/e9shx.

Indirect self-report scales are often used in hopes of improving measurement of explicit racial attitudes, yet none were

better at predicting implicit attitudes than a single item assessing overt preference for White versus Black people. These results suggest that while some people may alter responses due to social desirability concerns, the best way to measure explicit racial attitudes remains to ask about them directly.

Authors' Note

Declaration of Conflicting Interests

Funding

Supplemental Material

The supplemental material is available at https://osf.io/e9shx.

Notes

1. Ratings were completed before data collection. Originally, several items were removed because they assessed potentially outdated issues (e.g., school busing) or asked about hypothetical scenarios that were no longer hypothetical (e.g., a Black president). For the sake of having full scales, these items were later added back after coder ratings were completed. These items were Item 11 from attitudes toward Whites, Item 7 from modern racism, Item 5 from new racism, and Items 3 and 11 from racial attitudes. The thermometer items were originally excluded but later added to increase similarity with other Project Implicit data sets.
2. Scales with equal number of response options were averaged and standardized, except for two (prejudice index, subtle and blatant prejudice) with prespecified scoring. Scales with items containing a differing number of response options were standardized within items and then averaged, except for one (new racism) with prespecified scoring.

References

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations. *Psychological Bulletin*, *84*, 888–918.

Altemeyer, R. A. (1981). *Right-wing authoritarianism*. Winnipeg, Canada: University of Manitoba Press.

Alwin, D. F., & Krosnick, J. A. (1991). The reliability of survey attitude measurement the influence of question and respondent attributes. *Sociological Methods & Research*, *20*, 139–181.

Blincoe, S., & Harris, M. J. (2009). Prejudice reduction in white students: Comparing three conceptual approaches. *Journal of Diversity in Higher Education*, *2*, 232–242.

Bobo, L., & Kluegel, J. R. (1993). Opposition to race-targeting: Self-interest, stratification ideology, or racial attitudes? *American Sociological Review*, *58*, 443–464.

Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, *23*, 1933–1967.

Britt, T. W., Bonieci, K. A., Vescio, T. K., Biernat, M., & Brown, L. M. (1996). Intergroup anxiety: A person × situation approach. *Personality and Social Psychology Bulletin*, *22*, 1177–1188.

Cao, J., & Banaji, M. R. (2016). The base rate principle and the fairness principle in social judgment. *Proceedings of the National Academy of Sciences*, *113*, 7475–7480.

Davis, J. A., & Smith, T. W. (1991). *The NORC general social survey: A user's guide* (Vol. *1*). Thousand Oaks, CA: Sage.

De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 176–193). New York, NY: Guilford Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368.

Derous, E., Nguyen, H. H., & Ryan, A. M. (2009). Hiring discrimination against Arab minorities: Interactions between prejudice and job characteristics. *Human Performance*, *22*, 297–320.

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.

Duckitt, J., & Farre, B. (1994). Right-wing authoritarianism and political intolerance among whites in the future majority-rule South Africa. *The Journal of Social Psychology*, *134*, 735–741.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326.

Forscher, P. S., Cox, W. T., Graetz, N., & Devine, P. G. (2015). The motivation to express prejudice. *Journal of Personality and Social Psychology*, *109*, 791–812.

Friese, M., Hofmann, W., & Schmitt, M. (2009). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology*, *19*, 285–338.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731.

Goff, P. A., Steele, C. M., & Davies, P. G. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology*, *94*, 91–107.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

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*, 17–41.

Groom, V., Bailenson, J. N., & Nass, C. (2009). The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence*, *4*, 231–248.

Guinote, A., Willis, G. B., & Martellotta, C. (2010). Social power increases implicit prejudice. *Journal of Experimental Social Psychology*, *46*, 299–307.

Gulker, J. E., & Monteith, M. J. (2013). Intergroup boundaries and attitudes: The power of a single potent link. *Personality and Social Psychology Bulletin*, *39*, 943–955.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*, 23–34.

Henry, P. J., & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology*, *23*, 253–283.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385.

Hofmann, W., Gschwendner, T., Castelli, L., & Schmitt, M. (2008). Implicit and explicit attitudes and interracial interaction: The moderating role of situationally available control resources. *Group Processes & Intergroup Relations*, *11*, 69–87.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit—explicit consistency? *European Review of Social Psychology*, *16*, 335–390.

Huntsinger, J. R., & Smith, C. T. (2009). First thought, best thought: Positive mood maintains and negative mood degrades implicit-explicit attitude correspondence. *Personality and Social Psychology Bulletin*, *35*, 187–197.

Inzlicht, M., Gutsell, J. N., & Legault, L. (2012). Mimicry reduces racial prejudice. *Journal of Experimental Social Psychology*, *48*, 361–365.

Jacobson, C. K. (1985). Resistance to affirmative action: Self-Interest or racism? *Journal of Conflict Resolution*, *29*, 306–329.

Joy-Gaba, J. A., & Nosek, B. A. (2010). The surprisingly limited malleability of implicit racial evaluations. *Social Psychology*, *41*, 137–146.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893–905.

Kinder, D. R., & Sanders, L. M. (1996). *American politics and political economy. Divided by color: Racial politics and democratic ideals*. Chicago, IL: University of Chicago Press.

Kluegel, J. R., & Smith, E. R. (1986). *Beliefs without inequality: Americans' view of what is and what ought to be*. Hawthorne, NJ: Aldine de Gruyter.

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, *109*, 901–931.

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of anti-prejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, *22*, 1472–1477.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *140*, 1–55.

Lybarger, J. E., & Monteith, M. J. (2011). The effect of Obama saliency on individual-level racial bias: Silver bullet or smokescreen? *Journal of Experimental Social Psychology*, *47*, 647–652.

Maison, D., Greenwald, A. G., & Bruin, R. (2001). The implicit association test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin*, *32*, 1–9.

McClosky, H. (1958). Conservatism and personality. *American Political Science Review*, *52*, 27–45.

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, Discrimination, and Racism* (pp. 91–125). San Diego, CA: Academic Press.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565.

Nosek, B. A. (2007). Implicit–explicit relations. *Current Directions in Psychological Science*, *16*, 65–69.

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*, 101–115.

Nosek, B. A., Bar-Anan, Y., Sriram, N., Axt, J., & Greenwald, A. G. (2014). Understanding and using the brief implicit association test: Recommended scoring procedures. *PLoS One*, *9*, e110938.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). 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, NY: Psychology Press.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the implicit association test. *Experimental Psychology*, *54*, 14–29.

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*, 421–433.

Paulhus, D. L. (1984). Two-component models of social desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.

Payne, B. K., Krosnick, J. A., Pasek, J., Lelkes, Y., Akhtar, O., & Tompson, T. (2010). Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology*, *46*, 367–374.

Peffley, M., Hurwitz, J., & Sniderman, P. M. (1997). Racial stereotypes and whites' political views of blacks in the context of welfare and crime. *American Journal of Political Science*, *41*, 30–60.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, *25*, 57–75.

Phillips, S. T., & Ziller, R. C. (1997). Toward a theory and measure of the nature of nonprejudice. *Journal of Personality and Social Psychology*, *72*, 420–434.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Ranganath, K. A., & Nosek, B. A. (2008). Implicit attitude generalization occurs immediately: Explicit attitude generalization takes time. *Psychological Science*, *19*, 249–254.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.). (1999). *Measures of social psychological attitudes, Vol. 2. Measures of political attitudes*. San Diego, CA: Academic.

Rudman, L. A., & Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Processes & Intergroup Relations*, *5*, 133–150.

Saucier, D. A., & Miller, C. T. (2003). The persuasiveness of racial arguments as a subtle measure of racism. *Personality and Social Psychology Bulletin*, *29*, 1303–1315.

Sears, D. O. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53–84). New York, NY: Plenum.

Schmidt, K., & Axt, J. R. (2016). Implicit and explicit attitudes toward African Americans and Barack Obama did not substantively change during Obama's presidency. *Social Cognition*, *34*, 559–588.

Schmidt, K., & Nosek, B. A. (2010). Implicit (and explicit) racial attitudes barely changed during Barack Obama's presidential campaign and early presidency. *Journal of Experimental Social Psychology*, *46*, 308–314.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Sidanius, J., Pratto, F., Martin, M., & Stallworth, L. M. (1991). Consensual racism and career track: Some implications of social dominance theory. *Political Psychology*, *12*, 691–721.

Sidanius, J., Van Laar, C., Levin, S., & Sinclair, S. (2004). Ethnic enclaves and the dynamics of social identity on the college campus: The good, the bad, and the ugly. *Journal of Personality and Social Psychology*, *87*, 96–110.

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 7710–7715.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, *41*, 157–175.

Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial bias in perceptions of others' pain. *PLoS One*, *7*, e48546.

Uhlmann, E. L., Brescoll, V. L., & Machery, E. (2010). The motives underlying stereotype-based discrimination against members of stigmatized groups. *Social Justice Research*, *23*, 1–16.

Wald, K. D., Owen, D. E., & Hill, S. S. (1988). Churches as political communities. *American Political Science Review*, *82*, 531–548.

Author Biography

Jordan R. Axt is a postdoctoral research associate at the University of Virginia and Project Implicit. His research explores how people form and express intergroup bias in attitudes and behavior.