

The Black-White Malleability Gap in Implicit Racial Evaluations: A Nationally Representative
Study

Kevin Pinkston, PhD

University of Illinois at Chicago

Data collected by Time-sharing Experiments for the Social Sciences, NSF Grant 0818839,
Jeremy Freese and James Druckman, Principal Investigators

Abstract

This study replicates and extends the experimental design originated by Dasgupta and Greenwald (2001) who found a decrease in implicit pro-White biases after exposure to pictures of admired Black individuals. A nationally representative sample was analyzed comparing implicit pro-White biases among Black and White participants. Hypothesis 1 predicted a replication of previous research among White participants, and hypothesis 2, derived from the balanced identity theory, predicted an increased pro-Black bias among Blacks after exposure to admired Black individuals. Results provided partial support for hypothesis 1 and a lack of support for hypothesis 2. This is the first study to use a nationally representative sample to examine implicit pro-White biases. System Justification Theory was used to explain the malleability gap in Black and White pro-racial biases.

Keywords: implicit attitudes, iat, race, black identity, system justification theory, balanced identity theory

Introduction

Black Americans are a heavily stigmatized racial group in the U.S. Although negative stereotypes toward this group are widespread, Blacks have been shown to maintain a pro-Black bias when surveyed using explicit measures. The pro-Black bias does not persist, however, when measured using implicit or automatic measures. Researchers have conducted several studies examining Black implicit racial biases using the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and have found mixed results with an overall indication of a neutral implicit pro-Black bias (Nosek, Banaji, & Greenwald, 2002; Nosek, Smyth, Hansen, Linder, Ranganath, Smith, Olson, Cugh, Greenwald, & Banaji, 2007). Previous research suggests that implicit pro-White biases may be malleable upon exposure to positive and negative racial stimuli (Dasgupta & Greenwald 2001; Joy-Gaba & Nosek 2010; Rudman, Ashmore, & Gary, 2001). Data on Black implicit attitudes typically consist of samples of college students and individuals who have self-selected onto a website. The current study replicates and extends Dasgupta and Greenwald's (2001) experiment (also replicated by Joy-Gaba & Nosek, 2010) with the purposes of examining the malleability of implicit attitudes in a nationally representative sample and comparing the malleability of implicit pro-Black/White biases. Previous data on implicit pro-White biases have been largely limited to college student (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000) and Internet samples (Nosek et al 2002; Nosek et al. 2007; Sabin, Nosek, Greenwald, Rivara, 2009; Joy-Gaba & Nosek, 2010). The current experiment allows for greater generalizability for its findings.

In accordance with previous research on the malleability of implicit attitudes, I hypothesize that exposure to admired Black and disliked White exemplars will lead to a small decrease in the implicit pro-White bias while exposure to negative Black and positive White

exemplars will not affect implicit pro-White bias. The balanced identity theory (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellot, 2002) is used to hypothesize that Black participants will experience a similar exemplar effect as Whites resulting in an increased pro-Black implicit bias. Results partially supported hypothesis 1 with evidence for the replication of the Joy-Gaba and Nosek (2010) finding of a slight decrease in implicit pro-White bias among Whites after exposure to positive Black exemplars. This finding was only present however, when the data were not weighted to be nationally representative. The nationally representative data revealed no significant effect of the exposure condition on implicit pro-White biases. Results did not support the second hypothesis of a similar exemplar effect among Blacks. There is evidence of the malleability of implicit pro-White biases among Whites but not of the malleability of implicit pro-Black biases among Blacks. The system justification theory explains that the lack of malleability in Blacks' implicit pro-Black bias may be a result of an internalized tendency to maintain the status quo of a perceived lower social and economic status of Blacks compared to Whites in U.S. society. The remainder of this article will, first, define implicit attitudes and present research on Black implicit attitudes. Next, it will derive hypotheses of the exemplar effect on the implicit pro-White bias of Whites and the implicit pro-Black bias of Blacks. Then, it will describe the methods of data collection and the results of the study. Finally, theoretical propositions from the system justification theory will be used to explain the differential malleability of Black and White implicit pro-White biases as well as implications for future research.

Implicit Attitudes: Definition and Measurement

Let us first establish a clear definition of implicit attitudes. Greenwald and Banaji (1995) define implicit attitudes as beliefs developed from past experiences activated without conscious awareness that influence judgments and actions. Stereotypical beliefs associated with racial group members automatically trigger a set of associations related to that term. Automatic associations connected to a group act as implicit biases that influence judgments, perceptions, and actions toward a group. Implicit attitudes are distinct from explicit attitudes or conscious, self-reported attitudes (Nosek et al., 2002). Studies comparing implicit attitudes to explicit attitudes found correlations ranging between 0.27 and 0.56 (correlations varied depending on the topic) showing that implicit and explicit beliefs are only moderately related to each other (Fazio & Olsen 2003; Nosek et al. 2007).

The IAT is currently the dominant method used to measure implicit attitudes. The IAT is a computer-based categorization task in which participants rapidly categorize terms or pictures associated with two social categories such as White or Black as well as terms associated with two evaluative categories such as good or bad. The social and evaluative categories are combined, and participants' response times (how long it takes them to match the terms to the categories in milliseconds) are measured. Individuals may respond more quickly when terms associated with their group are paired with pleasant words and terms associated with the out-group are paired with unpleasant words. When this takes place, there is an implicit preference for the in-group. On the other hand, if individuals respond more quickly when in-group exemplars are paired with unpleasant words and out-group exemplars are paired with pleasant words this is a sign of an implicit preference for the out-group. Researchers found that Whites were quicker at accurately identifying a pair of concepts when Black exemplars were matched with unpleasant

words and White exemplars were matched with pleasant words (Greenwald, et al.1998).

Research from Dovidio and colleagues revealed behavioral consequences of implicit biases (Dovidio, Kawakami, & Gaertner, 2002). Greater implicit pro-White biases in Whites (in IATs comparing Blacks to Whites) predicted un-friendly behavior toward Black confederates. The IAT shows evidence of an overall implicit pro-White bias among Whites and of a relationship between a pro-White bias and behavior toward Blacks.

Arkes and Tetlock (2004) and Tetlock (2008) argue against the primary method of measurement of implicit attitudes, the IAT. They claim it does not measure implicit prejudice but rather learned stereotypes. Quillian (2008) replies that regardless of what these associations are called, they are distinctly different from explicit associations and they should be further researched because these attitudes lead to discrimination behaviors (see Lane, Banaji, Nosek, & Greenwald, 2007). Greenwald and others explain that knowledge of statistical differences is a causal factor in the development of implicit prejudice (Greenwald and Banaji 1995; Jost and Banaji 1994; Devine 1989). However, Quillian and Pager (2001) found evidence that knowledge of statistical differences is only a part of unconscious racism when assessing perceived crime in neighborhoods. Regardless of the specific source of these unconscious associations, discriminatory behavior occurs in relation to these associations and there is ample evidence in support of what Quillian (2008) calls unconscious racism (Dovidio et al., 2002). Similarly, sociologists have described this internalized prejudice as symbolic racism in which whites endorse American values of hard work, the Protestant ethic and use it as an explanation for racial/ethnic differences in achievement (Kinder & Sears, 1981; Krysan, 2000; Bobo & Fox, 2003). The IAT is a useful tool for decreasing the threat of social desirability effects in assessing group biases.

Background Research

Research examining Black implicit pro-Black biases provides mixed evidence of an in-group bias. Nosek and colleagues were among the first to measure Blacks' implicit attitudes. In a comparison of implicit and explicit attitudes among Blacks, Cohen's d effect sizes were calculated to estimate racial group biases (Nosek et al., 2002). There was a large explicit pro-Black bias ($d=-0.80$) but a slight implicit pro-White bias among Blacks ($d=0.16$). Nosek and colleagues (2007) yielded slightly different findings in a later assessment of Black implicit attitudes ($d=-0.05$). In this study, the overall pro-Black bias was calculated by averaging the adult Black/White IAT bias ($d=-0.05$) with two other racial IATs (skin tone differences, $d=0.22$; and racial differences in children, $d=-0.16$). After averaging all three IAT's, Blacks exhibited virtually no implicit pro-Black bias, or a miniscule pro-White bias ($d=0.01$). Researchers concluded that Blacks have no significant implicit pro-Black bias and note that Blacks are the only racial group to exhibit this lack of pro in-group racial bias.

The samples used to compile most data on implicit attitudes are not representative of the U.S. population. The project implicit demonstration web site is a major source for implicit attitudes data. The web site, launched in 1998, administers a variety of IATs including those on stereotypes, racial attitudes, and political preferences, to its visitors and collects data on their scores. The authors describe that the data collected from this website are not representative of the nation for two reasons. The first reason is self-selection. The individuals who complete IATs on this web database have chosen to do so. Individuals who choose this are likely to be educated and to have an interest in psychology and/or the topic. The second reason that the data are most likely not representative is the possibility of multiple submissions from one individual. The demonstration website has no control over the number of IATs participants complete or whether

or not participants repeat assessments (Nosek et al., 2002, Nosek et al., 2007). With the possibility of multiple IATs by the same individuals, people who take the test repeatedly may influence the data. The primary benefit of this sampling method is its ability to capture a large number of participants. The website has collected millions of IATs since its original launch. Greenwald, Nosek, and colleagues acknowledge these shortcomings in their studies. Nosek and colleagues have provided evidence using their web-based dataset that Blacks have no implicit pro-Black bias, yet their sample does not represent the nation.

Black implicit attitudes have largely been neglected in social psychological research. The current study aims to isolate Black implicit racial biases. Black implicit racial biases have significant behavioral implications which merit closer examination. There is evidence that Blacks with low implicit pro-Black biases are less willing to cooperate with their racial group members. Ashburn-Nardo and colleagues (Ashburn-Nardo & Monteith, 2003; Ashburn-Nardo, Monteith, Arthur, & Bain, 2007) provide evidence of the behavioral implications of this low pro-Black bias. In their 2003 study, 83 Black participants were told they would be completing an intellectually challenging task with a partner and were examined using the IAT. Pro-Black biases were positively correlated with a preference for Black work partner, even when controlling for explicit attitudes relating to partner preference.

Livingston (2002) reported two studies examining the variability in Black implicit in-group bias using the IAT. Participants who perceived greater negativity toward their group were more likely to report an implicit pro-White bias than participants who perceived less negativity toward their group. Participants who perceived the most negativity toward their group exhibited an implicit pro-White bias. These findings support the notion that implicit attitudes significantly correlate with perceptions of a group.

Previous research has also shown evidence that implicit attitudes can change after brief exposure to positive and negative exemplars. Dasgupta and Greenwald (2001) conducted an experiment in which exposure to pictures of positively evaluated Black exemplars and negatively evaluated White exemplars decreased the pro-White bias of their non-Black sample. In this experiment, participants were exposed to 10 admired black figures and 10 disliked white figures in one condition, and exposed to flowers and insects in a control condition. Their results showed that participants had significantly decreased implicit pro-White biases in the experimental condition compared to the control condition. The exemplar effect remained 24-hours after the experiment. Joy-Gaba and Nosek (2010) replicated this study in three experiments. They found an overall effect size that was noticeably smaller than that of Dasgupta and Greenwald (2001) (the effect size in the Joy-Gaba and Nosek (2010) was one tenth of the effect size in the Dasgupta and Greenwald (2001) experiment). Joy-Gaba and Nosek (2010) also found a significant difference between the positive Black exemplar condition and the control condition in one of their three experiments. Participants were sampled from a pool several other studies from the project implicit website. In another experiment, Joy-Gaba and Nosek (2010) found non-significant effect size differences between the two conditions with three different samples: 1) undergraduates in the laboratory; 2) undergraduates on the Internet; and 3) volunteers on the project implicit website (average effect size, $d=0.08$). One experiment conducted in the Internet sample (experiment 2a in Joy-Gaba & Nosek, 2010) yielded a significant positive effect of the admired exemplar condition ($d=0.17$). Researchers concluded that there is malleability in implicit pro-White biases, however, not as much as originally believed. Both of these studies excluded Blacks from their sample. Their rationale for this exclusion was that Blacks had already been shown to

have no significant implicit bias. The current study includes Blacks in the experiment in order to compare the exemplar effect in Blacks and Whites.

Hypotheses

The current study poses two questions derived from the research of Joy-Gaba and Nosek (2010) and Dasgupta and Greenwald (2001) experiments. First, does the exemplar effect observed in the two prior experiments on non-Black participants replicate in a nationally representative sample? Second, is there a differential exemplar effect across Blacks and Whites? Based on previous research on the limited malleability of implicit attitudes described above, I hypothesize that, in this nationally representative sample, exposure to admired Black exemplars and disliked White exemplars will result in an increased pro-Black bias among Blacks and a decreased pro-White bias among Whites. Also, exposure to disliked Black exemplars and admired White exemplars will not affect implicit pro-White/Black biases; the theoretical cause of this is not known.

The balanced identity theory is used to hypothesize an exemplar effect of an increased Black pro-Black bias. According to the balanced identity theory (Greenwald et al., 2002), implicit in-group biases are a function of one's self-esteem and group identity. The logic used to establish an in-group bias is, "if I am good and I am X then X is good". Rudman and Goodwin (2004) used the balanced identity theory to explain women's implicit pro-women bias. Women's high implicit self-esteem and high gender identity significantly interacted to predict their implicit pro-women bias. With regard to implicit pro-Black biases, Blacks who view admired Black exemplars may experience an increased Black identity as a result of being required to correctly identify the exemplar's racial group. Admired Black exemplars may prime Black participants to perceive themselves as apart of the racial group. A group association fulfills one of the two

criteria described by the balanced identity theory as necessary to foster a pro-group implicit bias (if I am X). An increased group identity may increase the implicit pro-Black bias given positive attitudes toward oneself, described by the balanced identity theory as the other necessary component necessary in fostering a pro-group implicit bias (if I am good). In feeling connected to the admired exemplars in the study, Blacks may display an increased average implicit pro-Black bias. In sum, Hypothesis 1 predicts a replication of the exemplar effect on White's implicit pro-White bias, and hypothesis 2 predicts an exemplar effect of an increased implicit pro-Black bias among Blacks.

Methods

Participants

Participants were recruited to this study through a custom research data collection service called GFK.¹ GFK recruited a pool of participants randomly across the United States from an address-based sample (ABS) from the U.S. Postal Service's Delivery Sequence file that covers about 97% of U.S. addresses. The organization sent recruitment letters to the households and then followed them up with phone calls. If the households did not have Internet capabilities, they were provided with a laptop with Internet access to participate in the study. Participants were placed on a panel to be called upon for participation in any GFK study. Panel members received incentives for their participation to increase panel loyalty and to decrease attrition within studies. For participants who were provided Internet services, the incentive was the hardware and free Internet service. Panel members who were not provided Internet services were incentivized via entrance into a rewards program in which they would receive points for participation in various panel studies. GFK distributed cash-equivalent checks to panel members every four to six

months based on the amount of points they have accumulated through their study participation.

Checks averaged between \$4 and \$6 per month.

[Table 1 about here]

Sample

The sample contained Black and White participants only, with a 20% oversample of Blacks. Table 1 provides a demographic comparison between the samples gathered by Dasgupta and Greenwald (2001), Nosek and colleagues (2007), Joy-Gaba and Nosek (2010), the present study, and the U.S. Census. Nosek and colleagues (Nosek et al., 2007) used data from the implicit website. As shown in table 1, the present sample and the census calculations both show 48% of the population being men. The current sample also has a similar percentage of individuals with a college degree in comparison to the census population estimates. The census population estimate for those with a bachelors degree or higher is 29% whereas the current weighted sample estimates 27% as having a bachelors degree or higher. Nosek and colleague's (2007) study has a lower percentage of men in their samples than the population and present studies. The median household income in the present sample is also comparable to the observed median household income with \$55,000 in the present sample, compared to \$53,046. This measure should be interpreted with caution however because the present study calculates household income by coding categorical data to midpoints. There were a total of 648 participants: 356 were White and 292 were Black. Of the 648 in the sample, 18 individuals were dropped because they did not consent to the study. Three cases were deleted because 10% of their trials took less than or equal to 300 milliseconds as explicated by the IAT D scoring algorithm (Greenwald, Nosek, & Banaji, 2003). The final sample analyzed had a total of 627 participants with 348 whites and 279 blacks.

Procedure

Procedures replicated an experiment originally conducted by Dasgupta and Greenwald (2001). The replication was extended to focus on racial differences in implicit in-group biases. A 3 (exemplar exposure: pro-Black, pro-White, and control) X 2 (participants' race: Black and White) design was used to explore the difference in the exemplar effect on Black and White implicit pro-racial bias. Individuals were randomly assigned into one of three conditions: pro-Black exemplar, pro-White exemplar, or the control condition. In the first condition, participants viewed admired Black and disliked White exemplars. In the second condition, participants viewed disliked Black and admired White exemplars. In the control condition, participants viewed neutral stimuli including insects and flowers. Participants were shown pictures of the exemplars under the belief that they were completing a "general knowledge" task to identify famous Americans. For example, in condition 1, participants viewed a picture of Martin Luther King, Jr. The picture was accompanied by a correct description, ("Leader of the Black Civil Rights movement in the 1960's") and an incorrect description ("Former Vice-President of the United States"). In the control condition, participants viewed a picture of a rose accompanied by a correct description, ("Rose") and an incorrect description ("Lily"). The task was to match the name to the description to the picture. Participants viewed the name of each exemplar again and categorized each person as Black or White.ⁱⁱ After completing the "general knowledge task" participants took the racial Attitudes IAT. Finally, explicit attitudes were measured using two thermometer scales. Participants were asked to rate how warm (favorable) they felt towards Blacks and Whites (on a scale of 0-100). Due to time restraints, the current study did not assess the participants IAT in a 24-hour follow-up as was done by Dasgupta and Greenwald (2001).

Exemplars

Participants viewed pictures of six famous Black and White individuals: three admired Blacks (Denzel Washington, Colin Powell, and Martin Luther King), three disliked Blacks (Mike Tyson, O.J. Simpson, and Marion Barry), three admired Whites (Tom Hanks, Peter Jennings, and John F. Kennedy), and three disliked Whites (Ted Kaczynski, Howard Stern, and Al Capone).ⁱⁱⁱ Racial exemplars and descriptions were taken from the list of exemplars used by Dasgupta and Greenwald (2001).^{iv} Each positive picture was followed with only positive response choices, and each negative picture with only negative choices so respondents would know that the person was either admired or disliked even if the respondent did not recognize the picture. The control condition used pictures of three flowers and three insects.

IATs

Participants completed a computer-administered Implicit Association Test (IAT) to assess racial bias (Greenwald et al., 1998). The IAT used four evaluative categories in sets of two on either corner of the screen (e.g. pleasant and black), and the associated words or pictures flash in the middle of the screen. Quicker associations with terms indicate a stronger association of that term with that concept. In accordance with the procedural design by Nosek, Greenwald, and Banaji (2007), the IAT consisted of seven blocks. The first block consisted of twenty trials that sort black and white terms using the keys “d” or “k” (e.g. black with d and white with k). In the second block participants sorted the evaluative terms (e.g. pleasant or unpleasant) for twenty trials. In the third block participants sorted the terms from all four categories in twenty trials (e.g. black with pleasant and white with unpleasant). In block four participants also sorted all four categories using the two keys but there were forty trials. Block five had forty trials that sorted black and white using the opposite keys as the previous blocks (e.g. black with “k” and white

with “d”). Block six sorted all four categories in the opposite way as blocks three and four such as black with bad and white with good in twenty trials. Block seven did the same thing as block six in forty trials. For each block the category labels (Black and White) appeared on the top left and right of the computer screen. When an individual made an error in categorizing the stimuli, a red X appeared to indicate the error and participants were told to correct their answer. The interval between the occurrence of one trial’s response and the presentation of the next stimulus was about 250 milliseconds. Half of the sample switched the order (swapping blocks 5-7 with blocks 1, 3 and 4). The IAT scores were calculated using the D scoring algorithm (Greenwald et al., 2003).^v In this sample, a positive D score represents a pro-White bias because respondents took less time to associate Black images with positive ideas than to associate Black images with positive ideas with positive ideas. D Scores have a possible range of -2 to +2 where D=0.15 represents a slight bias D=0.35 represents a moderate bias and D=0.65 represents a large bias (effect size estimates taken from project implicit). All trials over 10,000 milliseconds were deleted, as the IAT scoring regulates.

Analysis

[Table 1 about here]

In replication of the analysis conducted by Dasgupta and Greenwald (2001) and Joy-Gaba & Nosek (2010), ANOVA tests and *t*-tests compared the un-weighted pro-Black condition to the control condition, the pro-White condition to the control condition, and the pro-Black condition to the pro-White condition. In an expansion of the prior experiments, each condition was compared across race/ethnicity. Table 1 presents the demographic comparison between conditions and finds no significant differences between them. Exemplar effects on *explicit* attitudes were also analyzed using ANOVA tests and *t*-tests.

In a further extension of the previous experiments, sampling weights were added in Stata to the survey data to view the exemplar effect on a representative data set. Analyses used *t*-tests to compare Black and White implicit pro-racial biases across conditions with the weighted data. Regression analysis was also used to test and compare the exemplar effect across races while controlling for the demographic variables age, education, and gender. The exemplar condition variable was dummy coded for a comparison between the control condition and the pro-Black condition with the control condition as the omitted category (the pro-White condition was dropped from this analysis). The education variable was ordinal and consisted of four categories; one represented an individual with less than a high school education; two represented an individual with a high school diploma; three represented an individual with some college; and four represented an individual with a bachelor's degree or higher. Gender was dummy coded so that a one represented a man and zero represented a woman.

Results

[Figure 1 about here]

Results revealed a low correlation between implicit and explicit racial attitudes ($r=0.39$), and explicit attitudes were not different across conditions. Figure 1 compares the weighted IAT D scores across conditions and illustrates large differences in implicit pro-racial biases across races. Overall, Whites had a moderate implicit pro-White bias ($M=0.49$, $SE=0.02$) whereas Blacks had a slight pro-Black bias ($M=-0.13$, $SE=0.02$). Table 2 analyzed White participants' implicit pro-White bias across pro-Black, pro-White, and control conditions. There was no significant difference observed among Whites' pro-White implicit bias across all three conditions $F(2, 347)=2.14$ $p=0.12$. Individual comparisons made between experimental and control conditions revealed some marginal differences of note. Whites in the pro-Black condition

($M=0.42$, $SD=0.37$, 95% $CI[0.36, 0.48]$) displayed a marginally significant reduction in the average implicit in-group pro-White bias compared to those in the control condition ($M=0.51$, $SD=0.33$, 95% $CI[0.45, 0.56]$), ($t=1.82$, $p<0.1$). The Cohen's d effect size of the pro-Black exemplar condition ($d=0.23$, 95% $CI[-0.02, 0.49]$) is within the 95% confidence of the effect size ($d=0.17$) calculated in study 2a of Joy-Gaba and Nosek (2010). The evidence provides partial support for hypothesis 1 and is consistent with Joy-Gaba and Nosek's (2010) finding of a slight effect of pro-Black exemplars on the implicit pro-White bias.

Table 3 compared Blacks' implicit pro-Black bias in the pro-Black exemplar condition to the pro-White exemplar and to the control conditions. There was no significant difference observed in Blacks' pro-Black implicit bias across all conditions $F(2, 278)=0.5$ $p=0.61$. Blacks in the pro-Black condition ($M=-0.14$, $SD=0.04$, 95% $CI[-0.09, -0.32]$) did not exhibit a significantly different implicit pro-Black bias than those in the control condition ($M=-0.10$, $SD=0.04$, 95% $CI[-0.02, -0.17]$), ($t=0.74$, $p=0.46$) or those in the pro-White condition ($M=-0.09$, $SD=0.03$, 95% $CI[-0.03, -0.15]$), ($t=0.98$, $p=0.33$). Results show no evidence of an exemplar effect among Black participants and do not support hypothesis 2.

[Table 2 about here]

[Table 3 about here]

After conducting t -tests of the weighted samples comparing Black and White implicit pro-racial biases across conditions, I found no statistically significant difference between Whites in the pro-Black condition ($M=0.46$, $SE=0.04$, $CI[0.34, 0.54]$) and in the control condition ($M=0.51$, $SE=0.04$, $CI[0.44, 0.58]$), $t=0.98$, $p=0.33$). There was also no exemplar effect present among Blacks in the positive exposure condition ($M=-0.15$, $SE=0.04$, 95% $CI[-0.07, -0.22]$) compared to the control condition ($M=-0.12$, $SE=0.05$, $CI[-0.02, -0.21]$, $t=0.51$, $p=0.61$).

As shown in table 4, the weighted OLS regression models reveal no exemplar significant

effect on implicit attitudes. This suggests that implicit attitudes are not easily influenced by a small number of exemplars. They also reveal no significant interaction effect between race and exemplar exposure conditions. After dummy coding the race variable with Blacks as the omitted category, the coefficient reveals that Whites have a significantly stronger implicit pro-White bias than Blacks ($B=-0.62$, $SE=0.04$, $p<0.01$). This result is expected, and provides evidence of a more implicit pro-White bias among Whites than among Blacks. Analyses of control variables reveal a significant relationship between age and pro-White bias and gender and pro-Black bias. With every year increase in age, the pro-white bias appears to increase by a D score of approximately 0.003 ($B=0.002$, $SE=0.001$, $p<0.01$). This finding corroborates previous research on the effect of age on implicit attitudes (Nosek et al., 2007). Also, males had a slightly higher implicit pro-White bias than women ($B=0.08$, $SE=0.04$, $p<0.05$). After applying population weights to the survey data, analyses reveal demographic correlates with implicit pro-White biases but no evidence of an exemplar effect.

[Table 4 about here]

Discussion

The results of this study increase our overall understanding of the malleability of implicit pro-White biases. It provides partial evidence in support of hypothesis 1 with the replication of Joy-Gaba and Nosek's (2010) finding of an exemplar effect on Whites' implicit pro-White bias in a nationally representative sample but provides no evidence of an exemplar effect on Blacks' implicit pro-Black bias as stated in hypothesis 2. The present study provides a more representative account of the malleability of implicit pro-White biases than previous studies that can be generalized to the U.S. population. One limitation of the current study is its modest sample size. The current sample size lacks the statistical power of the previous IAT studies, and

may be a contributing factor to the lack of greater observed statistical significance between experimental conditions. Although this limitation exists, a Black-White gap in the malleability of implicit pro-White biases has been exposed. There is evidence of a pro-Black exemplar effect among White Americans, but no evidence of such an effect among Blacks. Future research should ask why the observed implicit pro-White malleability bias gap exists and how can it be closed.

Despite finding no support for the balanced identity theory explanation of an exemplar effect on the pro-Black bias, the study does not refute this theory. One explanation for the lack of support for the balanced identity theory is that Black exemplars may not affect the group identity and self-esteem necessary to foster an increased implicit pro-Black bias. There were no measures of implicit self-esteem or implicit group identity as was present in previous research (Rudman & Goodwin, 2004) on the balanced identity theory. Women were shown to be a unique disadvantaged group in that they exhibited positive implicit self-esteem, positive group identity, and an implicit pro-woman bias. These characteristics may not be easily elicited by exposure to Black and White exemplars.

Insight from the system justification theory (SJT) is useful in explaining the Black-White malleability gap in implicit pro-White biases (Jost & Banaji, 1994; Jost & Banaji, 2004). Blacks' perception of their group's status in society is internalized, and there is a desire to maintain this perception that supersedes any exemplar effect. The SJT posits that individuals seek to maintain the societal status quo by internalizing beliefs about their group affiliations based on its perceived status in society. As a consequence, minorities are said to develop a false consciousness, defined as holding beliefs that are contrary to a group's interests that perpetuate their disadvantaged status in society (Jost & Banaji, 1994). In the U.S. Blacks have internalized

negative stereotypes toward their group such as laziness and disproportionate poverty (Steele, 1997). The internalization of negative stereotypes is evidence of a false consciousness developed to justify the low status position of disadvantaged groups. According to hypotheses 2 of the SJT, a group may be stereotyped differently depending on whether it is perceived to be high or low status (Jost & Banaji, 2004). Blacks who perceive their group to be low status may not be affected by admired exemplars, as the exemplars likely cannot change the individual's perception of an entire group's status. Black exposure to admired exemplars may actually be perceived as a threat to the status quo and foster stereotypes to justify their perception of being a low status group, if originally present. Whites are not theorized to carry a false consciousness and therefore are more likely to exhibit malleability in pro-White implicit biases than Blacks. In understanding the predictors of the Black pro-Black bias, future research should explore the relationship between Black's (and other minority's) perceived negative stereotypes toward their group and their implicit pro-Black bias. Blacks' perception of their group status may positively predict their implicit pro-racial Black bias.

Although not explicitly tested in the study, the election of President Obama in 2008 may have affected Black and White implicit pro-White biases in different ways. Whites may have viewed the President as an ongoing positive exemplar leading to an increased implicit pro-White bias. Blacks viewing the President as an exemplar would not affect their implicit pro-Black bias. A Black President may serve to increase the implicit pro-Black bias only to the extent that Blacks perceived an increased overall social status in society. Little information is still known about the magnitude of the Black implicit pro-Black bias. Past research has found mixed evidence on the phenomenon with some evidence suggesting no implicit pro-Black bias or a slight pro-White bias (Nosek et al., 2007; Nosek et al., 2002) and some evidence suggesting a

slight pro-Black bias (Ashburn-Narudo et al., 2007). As previously stated, a lack of an implicit pro-Black bias in Blacks has an established relationship with a willingness to collaborate with Blacks on meaningful tasks (Ashburn-Nardo et al., 2007). Future research should further explore the magnitude of the pro-Black bias in a representative sample.

References

- Arkes, R., & Tetlock, P. (2004). Attributions of implicit prejudice, or would Jesse Jackson' fail' the implicit association test? *Psychological Inquiry*, 15(4), 257-278. doi: 10.1207/s15327965pli1504_01
- Ashburn-Nardo, L., Knowles M., & Monteith, M. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, 21(1), 61–87. doi: 10.1521/soco.21.1.61.21192
- Ashburn-Nardo, L., Monteith, M., Arthur, S., & Bain, A. (2007). Race and the psychological health of African Americans. *Group Processes & Intergroup Relations*, 10(4), 471-491. doi: 10.1177/1368430207081536
- Bobo, L., & Fox, C. (2003). Race, racism, and discrimination: bridging problems, methods, and theory in social psychological research. *Social Psychology Quarterly*, 66(4), 319-332.
- Dasgupta, N., & Greenwald, A. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81(5), 800-814. doi: 10.1037//0022-3514.81.5.800
- Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5–18. doi: 10.1037/0022-3514.56.1.5
- Dovidio, J., Kawakami, K., & Gaertner, S. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62–68. doi: 10.1037//0022-3514.82.1.62
- Fazio, R., & Olson, M. (2003). Implicit measures in social cognition: their meaning and use. *Annual Review of Psychology*, 54(1), 297–327. doi:

10.1146/annurev.psych.54.101601.145225

- Greenwald, A., & Banaji, M. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27. doi: 10.1037/0033-295X.102.1.4
- Greenwald, A., Banaji, M., Rudman, L., Farnham, S., Nosek, B., & Mellott, D. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109(1), 3–25. doi: 10.1037/0033-295X.109.1.3
- Greenwald, A., McGhee, D., & Schwartz, J. (1998). Measuring individual differences in I implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, A., Nosek, B., & Banaji, M. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197. doi: 10.1037/0022-3514.85.2.197
- Jost, J. & Banaji, M. (1994). The role of stereotyping in system–justification and the production of false consciousness. *British Journal of Social Psychology*, 33(1), 1-27. doi: 10.1111/j.2044-8309.1994.tb01008
- Jost, J., & Banaji, M. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*. 25 (6): 881 919. doi: 10.1111/j.1467-9221.2004.00402
- Joy-Gaba, J., & Nosek, B. (2010). The surprisingly limited malleability of implicit racial evaluations. *Social Psychology*, 41(3), 137-146. doi: 10.1027/1864-9335/a000020
- Kawakami, K., Dovidio, J., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78(5), 871-888. doi:

10.1037/0022-3514.78.5.871

Kinder, D., & Sears, D. (1981). Prejudice and politics: symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, 40(3), 414-431. doi:

10.1037/0022-3514.40.3.414

Krysan, M. (2000). Prejudice, politics, and public opinion: understanding the sources of racial policy attitudes. *Annual Review of Sociology*, 26(1), 135-168. doi:

10.1146/annurev.soc.26.1.135

Lane, K., Banaji, M., Nosek, B., & Greenwald, A. (2007). Understanding and using the implicit association test: IV (pp. 59–102). In B. Wittenbrink and N. Schwarz *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.

Livingston, R. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, 38(4), 405-413. doi: 10.1016/S0022-1031(02)00002-1

Nosek, B., A., Banaji, M., & Greenwald, A. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics*, 6(1), 101-15. doi: 10.1037/1089-2699.6.1.101

Nosek, B., Smyth, F., Hansen, J., Devos, T., Lindner, N., Ranganath, K., Smith, C., Olson, K., Cugh, D., Greenwald, A., & Banaji, M. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, 18(1), 36-88.

doi:10.1080/10463280701489053

Nosek, B., Greenwald, A., & Banaji, M. (2007). The implicit association test at age 7: A methodological and conceptual review (pp. 265–292). In John A. Bargh *Automatic Processes in Social Thinking and Behavior*. Psychology Press.

Quillian, L. (2008). Does unconscious racism exist? *Social Psychology Quarterly*, 71(1), 6-11.

doi: 10.1177/019027250807100103

Quillian, L., & Pager, D. (2001). Black neighbors, higher crime? The role of racial stereotypes in evaluations of neighborhood crime. *American Journal of Sociology*.

107(3), 717-767. doi: 10.1086/338938

Rudman, Ashmore, R. & Gary, M. (2001). Unlearning automatic biases: the malleability of implicit prejudice and stereotypes. *Journal of personality and social psychology*, 81(5),

856. doi: 10.1037/0022-3514.81.5.856

Rudman, L. & Goodwin, S. (2004). Gender differences in automatic In-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*,

87(4), 494-509. doi:10.1037/0022-3514.87.4.494

Sabin, J., Nosek, B., Greenwald, A., & Rivara, F. (2009). Physicians implicit and explicit attitudes about race by MD race, ethnicity, and gender. *Journal of Health Care for the*

Poor and Underserved, 20(3), 896-913. .doi: 10.1037/0022-3514.87.4.494

Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and

performance. *American psychologist*, 52(6), 613. doi: 10.1037/0003-066X.52.6.613

Tables

Table 1. Demographic Comparison Across Samples and Population

	Nosek (2007)	J&N (2010)	D&G (2001)	Present Sample	U.S. Census
Age	26 (12)	28.8	N/A	47 (0.78)	37
Bachelors degree or higher	61%	0%	0%	27%	29%
Percent Male	38%	35%	48%	46%	46%
Median Income	N/A	N/A	N/A	55,000	53,046
Percent Black	6.70%	0%	0%	43%	13%
Percent White	72%	75%	65%	57%	78%
N	2,758,490	944	48	627	316,128,839

Note 1: Standard errors in parentheses where known.

Note 2: The Dasgupta and Greenwald (2001) sample consisted of college students only.

Note 3: Fifteen percent of Nosek (2007) sample is from non-U.S. nations.

Note 4: Joy-Gaba and Nosek (2010) used multiple samples in this study. The numbers above represent the sample taken from their internet sample in experiment two.

Note 5: the census' measure of individuals with a bachelors degree represent individuals 25 years of age and older.

Note 6: Census estimates are based on projected population estimates for 2013.

Note 7: The “present sample” values were calculated after weighting data to be representative of the population.

Table 2. White Implicit Pro-Racial Bias Comparison Across Conditions

Condition	Control	Pro-White Exemplar	Pro-Black Exemplar
Age	51 (17)	50 (17)	49 (17)
Education	2.88 (0.97)	2.91 (0.94)	2.86 (1.00)
Percent Men	43%	44%	45%
Implicit Pro-White Bias	0.51 (0.33)	0.49 (0.27)	0.42 (0.37)+
N	116	110	122

Note 1: Standard deviations are in parentheses.

Note 2: + $p < 0.1$ two-tailed t -test comparing mean to the control condition.

Note 3: A positive implicit pro-White score represents a pro-White implicit bias.

Table 3. Black Implicit Pro-Racial Bias Comparison Across Conditions

Condition	Control	Pro-White Exemplar	Pro-Black Exemplar
Age	47 (16)	49 (15)	48 (15)
Education	2.71(0.93)	2.69 (0.92)	2.78 (0.89)
Percent male	41%	48%	43%
Implicit Pro-Black Bias	-0.10 (0.40)	-0.09 (0.31)	-0.14 (0.37)
N	92	98	89

Note 1: Standard deviations are in parentheses.

Note 2: A negative implicit pro-Black bias score represents a pro-Black bias.

Table 4. OLS Comparison of Pro-Black and Control Condition with Controls N=418

	Model 1	Model 2	Model 3
Pro-Black Exemplar	-0.04 (0.04)	-0.03 (0.06)	-0.05 (0.05)
White	0.62*** (0.04)	0.63*** (0.06)	0.61*** (0.06)
White*Pro-Black Exemplar		-0.02 (0.08)	
Education			-0.01 (0.002)
Age			0.002** (0.001)
Men			0.08* (0.04)
Constant	-0.11** (0.04)	-.012** (0.05)	-0.12 (0.11)
R-squared	0.44	0.44	0.45

Note 1: Standard errors in parentheses.

Note 2: + p<0.1, * p<0.05, **p<0.01, ***p<0.01.

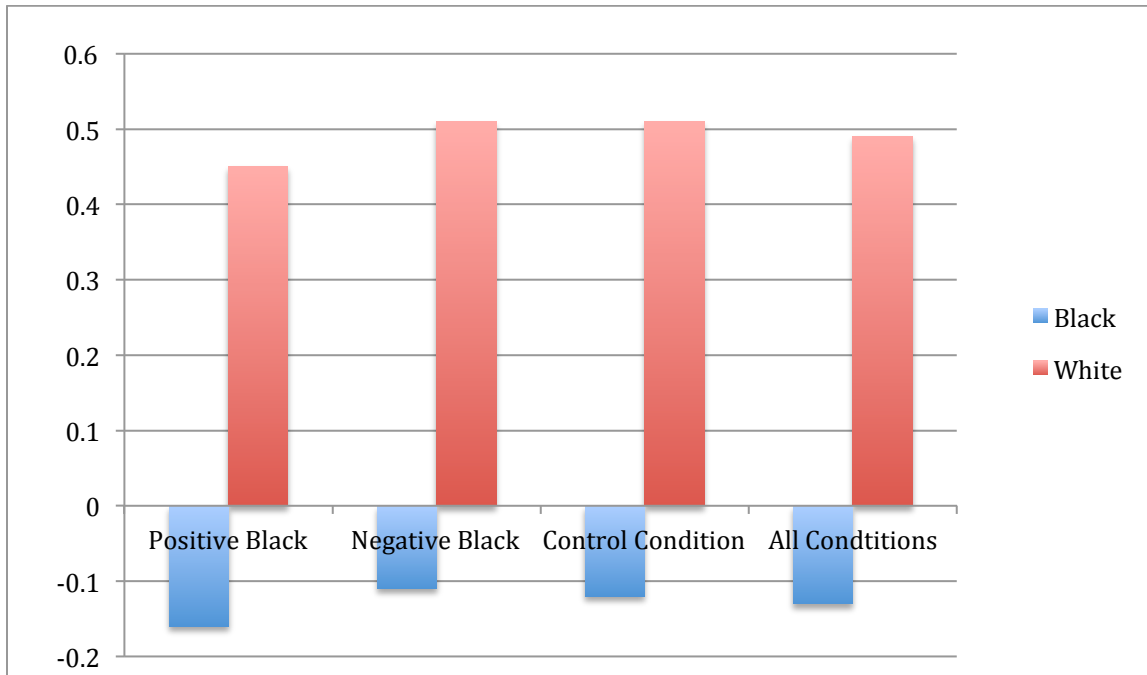
Note 3: Data were weighted to represent the population.

Note 4: A positive number represents an increase in implicit pro-White bias.

Figures

Figure 1. Mean D scores for the racial IAT of blacks and whites across experimental conditions

N=627



Note 1: The Y-axis represents IAT D scores.

Note 2: A positive D score represents a pro-White implicit bias.

Note 3: Data were weighted to represent the population.

Appendix A

Table A.1. Racial Exemplar Descriptions

Name	True Description	False Description
Martin Luther King	Leader of the Black Civil Rights movement in the 1960s.	Former Vice President of the United States.
Colin Powell	Former Chairman, Joint Chiefs of Staff for the U.S. Department of Defense.	U.S. ambassador to the United Nations.
Denzel Washington	Famous actor who played the leading role in the recent movie "Flight".	Famous golf champion.
Al Capone	American gangster who terrorized Chicago in the 1920s.	Leader of an anti- government militia.
Ted Kaczynski	The Unabomber who injured and killed using letter bombs.	Convicted pedophile.
Howard Stern	Notoriously offensive radio talk show host.	Accused of embezzlement.
O.J. Simpson	Charged with the brutal killing of his wife and friend	Embezzled millions of tax payers' money.
Mike Tyson	Suspended from boxing for biting off his opponent's ear.	Convicted pedophile.
Marion Barry	Former Mayor of Washington D.C. accused of buying and using cocaine.	Mafia member.
Tom Hanks	Received two Oscars for his acting roles in 'Philadelphia' and 'Forest Gump'.	Famous American Tennis Player
John F. Kennedy	Famous American President assassinated in Dallas.	Former Vice-President of the United States
Peter Jennings	News anchor for ABC	U.S. Ambassador to the United Nations

ⁱThe GFK acronym represents a German phrase translated into English as Society for Consumer Research.

ⁱⁱ Dasgupta and Greenwald (2001) displayed the exemplars four total times: twice before the first IAT block of 40 trials, and twice after the first block of 40 trials. The current study only displayed the exemplars twice before the IAT began. Joy-Gaba and Nosek (2010) also only displayed the exemplars twice before the IAT began.

ⁱⁱⁱ Dasgupta and Greenwald (2001) and Joy-Gaba and Nosek (2010) used 10 racial exemplars in their experiments. Due to the financial constraints of the TESS grant, the number of racial exemplars was limited to six exemplars per condition as opposed to 10.

^{iv} Appendix A provides descriptions of all the exemplars used in the experiment.

^v The IAT scores were calculated using the following eight step D scoring algorithm: “(1) Use data from blocks 3, 4, 6, and 7; (2) eliminate trials with latencies > 10,000 milliseconds; (3) eliminate subjects for whom 10 percent of trials have latencies < 300 milliseconds; (4) compute one standard deviation for all trials in blocks 3 and 6, and another standard deviation for all trials in blocks 4 and 7; (5) compute two difference scores (one between 3 and 6 and the other between 4 and 7), subtracting what is intended to represent the high (positive) end of the measure from the block containing associations representing the low end; (6) divide each difference score by its associated standard deviation from step 4; and (7) average the two quotients from step 6” (Greenwald et al., 2003). Trials that contained errors received a 600-millisecond penalty.