www.sciencemag.org/cgi/content/full/326/5960/1711/DC1

Supporting Online Material for

The Subtle Transmission of Race Bias via Televised Nonverbal Behavior

Max Weisbuch,* Kristin Pauker, Nalini Ambady

*To whom correspondence should be addressed. E-mail: max.weisbuch@tufts.edu

Published 18 December 2009, *Science* **326**, 1711 (2009)
DOI: 10.1126/science.1178358

This PDF file includes

SUPPORTING ONLINE MATERIAL

Study 1

Materials and Methods

Selection of Television Clips

A final sample of 30 characters (15 white, 15 black) was selected from 11 television shows and 9 short, silent clips were selected for each character. Please see Table S1 for a list of these characters and shows. The selection process began with television programs. We aimed to sample a variety of popular television programs in order to broadly estimate exposure to nonverbal race bias. To permit the estimation of bias in a fashion that could be generalized beyond a single day, week, or season, the sample was further restricted to shows with recurring themes and characters (television content otherwise varies widely on a daily basis). Scripted television shows and long-running reality shows fit criteria of popularity and recurrent themes/characters and these shows were the focus of Study 1. To control for any potential confound at the level of television show (e.g., white characters coming from more light-hearted shows), white and black characters were status-matched *within*, rather than between shows. Of the television shows that included both white and black recurring characters, many included a built-in confound in which black characters were much less central to the show's theme or had much lower job status. These programs were excluded. The remaining 11 programs were selected for Study 1 and for each black character in a program a white character with corresponding status was selected.

The authors identified the race of the characters *a priori*. To confirm that the characters were perceived to be of the intended race, undergraduate and graduate students were provided with partial course credit in exchange for identifying the race of the target characters. In one such pretest, participants (12 white, 2 Asian, 2 black, self-identified) viewed a still image of each character and identified their race as either white or black. These participants achieved 100% agreement with each other and with the authors' *a priori* designations. In a second pretest, 12 undergraduate participants (8 white, 1 Latino, 1 Asian, 2 Black) were asked to identify the race of characters from the character's name alone. These participants were asked to only attempt to identify race (white or black) for characters with whom they were familiar. These participants also achieved 100% agreement with each other and with the authors' *a priori* designations. In short, the characters were perceived to be the race designated by the authors.

To confirm that the characters were indeed matched for status, 15 undergraduate students (11 white, 2 black, 2 multi-racial, self-identified) rated the status associated with each character's job. Specifically, the authors referred to the website for each show to determine the job title of each character. Teenage target characters (who did not have jobs) were described with their main role (e.g., cheerleader, quarterback, fraternity pledge). Each job (or role) title was then rated by participants on Likert-type scales (from

0, extremely low status, to 6, extremely high status). These participants were consistent in their status ratings ($\alpha = .97$); thus, we computed a "status" score for each character from the average across the judges. As expected, the status of black characters ($M = 3.88$) was approximately equal to the status of white characters ($M = 3.84$), $t\ (28) = .07$, $p = .94$.

For each show, an effort was made to select 3 episodes appearing during the winter of 2006. In some cases the episodes purchased (via iTunes®) did not include enough clips for each character. Hence, while 3 episodes were utilized to select video clips for most shows, it was necessary to use up to 5 episodes for some of the shows. In general, the selection of clips comes from the same period (Winter 2006) across shows (for access to these video clips, please click on the appropriate link located at http://ase.tufts.edu/psychology/ambady/materials/weisbuch_et_al.htm).

A systematic and unbiased *a priori* procedure was utilized for selecting clips. Clips were only selected if they included 10-seconds of uninterrupted social interaction, included both characters in a single frame for at least one second, and occurred in the first, middle, or last 5 minutes of an episode. Moreover, only the featured character was non-white (i.e., the "expressers" were always white). In the event that multiple clips met these criteria, the first clip from each segment (early, middle, late) was always selected. Since impressions derived from 10-15 seconds of nonverbal exposure are as accurate as those derived from longer time periods (*S1*), clips were limited to 15 seconds or less to provide a measure of standardization across characters. There were a total of 265 clips (5 clips could not be located without sacrificing selection criteria).

Clip Ratings

Twenty-three white undergraduate students (12 female) were recruited via an advertisement that solicited students who rarely or never watched television. Students who replied to the advertisement were emailed a survey that required them to identify their race and gender and which of the 11 shows they had or had not watched. Those students who reported that they had not watched any of the shows were invited to participate at a $10 hourly rate (to a maximum of $50). These students rated the extent to which the featured character was "liked" by and "treated positively by" the other characters in that clip. To ensure that these "judges" were not influenced by the character's race or appearance and to ensure that ratings were based on nonverbal content, the ten-second clips for each character were digitally edited prior to rating. Specifically, the featured character within each clip was digitally edited out of the scene with Adobe Premiere® software by eliminating those parts of the screen that included the character. This ensured that the race or appearance of the featured character did not influence ratings of nonverbal responses *to* that character. This important design element discouraged demand characteristics that might otherwise result from the race of the featured character (who could not be seen).

The audio track was also digitally removed from each clip to ensure that only nonverbal information was available to judges. These 265 clips were presented in a random order to groups of two to five judges. These judges rated "how much the visible

characters liked or disliked the 'unseen' character" (from -3, strongly disliked to +3, strongly liked) and "the positivity of the interaction" (from -3, extremely negative, to +3, extremely positive).

Judges agreed in their ratings of liking ($\alpha = .85$) and positivity ($\alpha = .86$). Each character was thus assigned a liking score and a positivity score based on across-judge averages. These two scores were highly correlated, $r(28) = .87$, so their average was derived for each character to index the degree to which that featured character elicited a favorable nonverbal response (see Table S1).

To assess the existence of systematic dispositional differences between white and black characters, 17 white judges (12 female) were recruited via an advertisement soliciting students who regularly watched all or most of the sampled programs. Students who replied to the advertisement were emailed a survey that required them to identify their race and gender and which of the 11 shows they did or did not regularly watch. Judges who regularly watched no less than 8 of these shows were selected to rate the attractiveness and dispositions of featured characters and were paid for doing so. These judges were given the name and a picture of each character and were asked to rate each character on attractiveness, sociability, kindness, and intelligence (on scales ranging from 1, not at all, to 7, extremely).

To assess verbal content of the clips, an additional group of 13 white judges (7 female) was recruited via an advertisement for a study described as "rating people." Because specific television characters were not identifiable from transcripts, these student judges were not screened for television expertise but were screened for self-reported race and gender and were paid for their participation. These students rated the transcribed verbal content of each of the 265 clips described above and exhibited satisfactory inter-judge agreement with respect to liking ($\alpha = .86$). Each character was thus assigned a verbal liking score (Table S1).

Additional Statistics

As detailed in the main text, white characters elicited more favorable nonverbal responses than did black characters; no race differences emerged for favorable verbal responses, perceived attractiveness, sociability, kindness, or intelligence. Descriptive analysis of the 11 television shows provides a nice illustration of the nonverbal race bias exhibited on television. Weighting the nonverbal favoritism scores by race (-1 for black) and averaging these weighted scores within shows demonstrates that 9 of the 11 shows exhibited a pro-White nonverbal bias. Moreover, the median magnitude of pro-Black bias in the two "pro-Black" shows (*Median* = -.045) was only about half of the median magnitude of pro-White bias in the nine "pro-White" shows (*Median* = .086). These data illustrate that the existence of nonverbal race bias is not driven by a single show—it was a predominant theme of the 11 shows we sampled (for correlations among character ratings, see Table S2).

Study 2

Materials and Methods

Participants

Fifty-three white undergraduate students (37 female) were recruited to participate in exchange for partial course credit in an introductory psychology course. Information about race and gender were obtained via a "prescreening" questionnaire session completed by all students at the beginning of the semester.

Procedure

In individual cubicles, participants completed a series of short studies, the first two of which constituted this Study. The ostensible purpose of participants' first task was to collect data on students' television viewing preferences; as such, their only task was to indicate whether or not they watched each of 11 television shows (the shows from Study 1). In fact, this task allowed us to compute our predictor variable (exposure to televised nonverbal bias). Once participants had completed the simple task of listing their favorite shows, they were instructed on how to complete the second study, which was actually the outcome variable for Study 2. This was a standard implicit association test (IAT; *S2*) in which participants categorized faces as white or black and stimulus words (e.g., war, love, pleasant) as "positive" or "negative" words. Stimuli included 10 white male faces and 10 black male faces and 10 good and 10 bad words. In this study and all subsequent studies, information about the race and gender of the photographed individuals was provided in the original photograph collections (please contact authors for these images). All pictures were placed against a standardized grey background and re-sized to 300x450 pixels. For each trial, stimuli appeared one at a time in the center of the computer screen, and participants used the ''e'' and ''i'' keys on the computer keyboard to classify items as quickly and accurately as possible into the corresponding categories identified on the left or right side of the screen. Critical trials included congruent blocks where white/positive and black/negative shared response keys (block w-p) and incongruent blocks where white/negative and black/positive shared response keys (block b-p). The order of these critical blocks and response key mappings were counterbalanced between-subjects.

In typical studies, responses are fast when highly associated concepts, categories, and attributes share the same key (congruent blocks), and slow when weakly associated or inconsistent concepts, categories, and attributes share the same response key (incongruent blocks). Thus, we expected that (White) participants would strongly associate white with "positive" and black with "negative" but this pattern should vary with exposure to televised nonverbal bias.

Additional Statistics

To index media exposure to nonverbal race bias, the shows watched regularly by each participant were converted into by-show nonverbal bias scores. For each show, the nonverbal bias score was obtained by weighting favorable nonverbal response scores of

black and white characters by -1 and 1, respectively. For example, CSI would have a nonverbal bias score of .09, reflecting the average of weighted scores for Gil (+.227) and Warrick (-.055; see Table S10). Hence, shows with nonverbal bias scores greater than 0 exhibited relatively pro-white nonverbal bias whereas shows with scores less than 0 exhibited relatively pro-black nonverbal bias. The nonverbal bias scores for each show watched by a participant were then averaged to derive that participant's exposure score. For example, if a participant watched CSI and House the relevant nonverbal bias scores would be .09 and .06 (see Table S1). This participant would have an exposure score of .07. In Study 2 the average exposure score was .15 and scores ranged from -.01 to .60.

Race associations measured by the IAT were calculated according to the recommended *D* scoring algorithm (*S3;* see Table S3). Response latencies were calculated from the onset of the trial until a correct response was made, and latencies less than 300 ms or greater than 10,000 ms were removed. A positive IAT score indicates a stronger association between white faces and "positive", whereas a negative IAT score indicates a stronger association between black faces and "positive". As detailed in the main text, more exposure to nonverbal race bias was associated with greater IAT scores.

We also calculated by-show nonverbal bias scores for variables other than race. The nonverbal favoritism score for each character was multiplied by that character's perceived attractiveness score and these scores were averaged within show. For example, on CSI, the appropriate calculation would include Gil (.227 * 3.81) and Warrick (.055 * 5.25) so the "attractiveness bias" score would be .58 (the average of the two products). These scores index the extent to which each show depicted more attractive characters eliciting more positive nonverbal behavior than less attractive characters. We summed these scores across the shows watched by each Study 2 participant such that each participant was assigned an "exposure to nonverbal attractiveness bias score." These same procedures were utilized to compute exposure to nonverbal sociability, kindness, and intelligence, biases for each participant. As detailed in the main text, there was no significant correlation between race associations and exposure to nonverbal bias in favor of attractive people, sociable people, kind people, or intelligent people. And verbal race bias scores (calculated in the same manner as nonverbal race bias but with transcript ratings; M = .19, *range* = -.08 to 1.00) failed to correlate with IAT scores. Finally, controlling for these other biases failed to reduce the relationship between exposure to nonverbal bias and IAT scores (see main text).

Studies 3a and 3b

Materials and Methods

Participants

Sixty-two white participants (33 female; Study 3a) and thirty-five white participants (24 female; Study 3b) were recruited via advertisements for paid studies. As part of their response to the advertisement, participants identified their race, gender, and other demographic characteristics.

Materials

Study 3a video clips. Two sets of silent (nonverbal) video clips were constructed from the clips used in Study 1. Unlike Study 1, however, the featured character was *not* cropped out of these clips. Pro-white clips depicted the white characters eliciting positive nonverbal behavior the black characters eliciting negative nonverbal behavior (N = 58 clips). Pro-black clips depicted the black characters eliciting positive nonverbal behavior and the white characters eliciting negative nonverbal behavior (N = 60 clips). The same featured characters appeared in both conditions. For example, each white character was shown eliciting relatively positive nonverbal behavior in the pro-white clips and relatively negative nonverbal behavior in the pro-black clips. Each of the characters from Table S1 contributed clips to both conditions of Study 3a. By using the same characters in the two conditions, we ensured that differences between the conditions could not be attributed to differences in the characters but rather to the nonverbal behavior elicited by these characters. Indeed, the video ratings from Study 1 confirm that pro-white clips depicted white characters (M = 1.02) eliciting positive nonverbal behavior relative to black characters (M = -.86), $t(56) = 20.24$, $p < .001$, whereas the opposite was true for the pro-black set (Ms = -.70 and .69, respectively), $t(58) = 14.00$, $p < .001$.

Study 3b video clips. The purpose of Study 3b was to examine convergent validity with Study 3a. Specifically, new sets of pro-white and pro-black video clips were compiled for use in Study 3b. These clips were selected from a variety of television shows including but not limited to those used in Study 1 or Study 3a. New shows included Lost, the Real World, and The Closer; new characters were selected from House, CSI: Miami, and Heroes; and new clips from existing characters were added. These new characters and clips were supplemented with clips from Study 1 to produce a sample of video clips that was purposefully selected on the basis of clearly communicated nonverbal behavior (see Table S4).

To ensure that the clips differed in the intended fashion raters were recruited to judge the clips. For purposes of pretesting (but not the main study), we again cropped out the featured characters (and removed audio content) and recruited 17 white undergraduate students (13 females) to rate the extent to which these "unseen" characters were (a) liked by other characters and (b) involved in a positive interaction (on -3 to +3 scales). These raters agreed in their judgments (α's > .94); liking and positivity scores were consequently averaged across judges. These two scores were highly correlated [$r(28) = .92$] and were averaged to create a nonverbal favorability index. These ratings confirm that clips in the pro-white set depicted white characters (M = .64) eliciting positive nonverbal behavior relative to black characters (M = -1.11), $t(54) = 8.39$, $p < .001$, whereas the opposite was true for the pro-black set (Ms = -.94 and 1.07, respectively), $t(53) = 8.66$, $p < .001$.

Although many of the same characters appeared in the pro-white and pro-black conditions (see Table S4), it was difficult to locate characters that could each supply both several positive and several negative new clips. In cases where one character supplied positive clips and another supplied negative clips, an effort was made at matching those

characters on attractiveness, sociability, kindness, intelligence, and character role. To confirm that this effort was successful, the same 13 judges who rated attractiveness, sociability, kindness, and intelligence for Study 1 made those same ratings for the characters in Study 3b (including characters appearing in both conditions). As expected, the white characters in the pro-white and pro-black conditions exhibited similar levels of attractiveness (*Ms* = 5.45, 5.55), sociability (*Ms* = 4.46, 4.38), kindness (*M*s = 4.69, 4.46), and intelligence (*Ms* = 5.37, 5.37), *t*s (14) < .54, *p*s > .59. Similarly, the black characters in the pro-white and pro-black conditions exhibited similar levels of attractiveness (*Ms* = 4.61, 4.48), sociability (*Ms* = 5.00, 4.75), kindness (*M*s = 4.76, 4.58), and intelligence (*Ms* = 5.41, 5.49), *t*s (14) < .75, *p*s > .46.

Pilot Study

To ensure that the pattern of nonverbal behavior was subtle and difficult to consciously identify, we recruited 22 white paid participants (14 female) via an advertisement. These participants were randomly assigned to view either the pro-white or pro-black clips (randomly assigned) under instructions to try to identify a hidden pattern across the clips. After viewing these silent clips, participants were informed that the hidden pattern regarded race and nonverbal behavior. They were then given a two-choice question, only one of which was correct: (a) black people were treated better than white people, (b) white people were treated better than black people. As described in the main text, 5 of 11 participants in each condition selected choice "a", and this rate was not different from chance. Note that these participants were perhaps more vigilant than most television viewers (and participants in Studies 3a and 3b) in that the goal of these participants was to discover the hidden pattern. The fact that these pilot participants were nonetheless unable to identify this pattern suggests that it is indeed subtle and difficult to consciously identify. Hence, any influence of these clips should be considered subtle and perhaps nonconscious.

Procedure

In individual cubicles, participants were randomly assigned to watch either the pro-black or pro-white clip sets and then completed the same race IAT from Study 2. In the exposure phase, participants were informed that we were interested in impressions of silent video clips. Participants rated the perceptual clarity of each clip (these clips were presented in a different random order for each participant). After viewing these clips, participants were instructed to begin an ostensibly unrelated study regarding cognitive responses to faces and words. They then completed this race IAT, were thanked, paid, and debriefed (see Table S3 for calculation of IAT scores).

Additional Statistics

A 2 (experimental condition) by 2 (gender) independent-groups ANOVA on Study 3a IAT scores revealed only a main effect of experimental condition whereby participants in the pro-white condition exhibited higher IAT scores (*M* = .70) than participants in the pro-black condition (*M* = .43), $F(1,58) = 3.91$, $p = .05$, $rpb = .25$. There was no effect of gender [$F(1,58) = .83$, $p = .37$] nor an interaction [$F(1,58) = .36$, *p*

= .55]. For Study 3b, a 2 (experimental condition) by 2 (gender) independent-groups ANOVA revealed two main effects. Males exhibited higher IAT scores (M = .73) than did females (M = .47), $F(1,31) = 4.92$, p = .03, r(pb) = .37. More importantly, this analysis revealed the expected main effect of experimental condition, $F(1,31) = 4.75$, p = .04, r(pb) = .36, whereby participants in the pro-white exposure condition exhibited higher IAT scores (M = .70) than those in the pro-black exposure condition (M = .43). No significant interaction emerged, $F(1,31) = .004$, p = .95.

Study 4

Materials and Methods

Participants

Fifty-six white participants (35 female) were recruited via an advertisement for a paid study. As part of their response to the advertisement, participants identified their race, gender, and other demographic characteristics.

Materials

Video clips. For the pro-white and pro-black conditions, the video clips were identical to Study 3. As described in the main text (Study 3b), the pro-white condition contained clips in which white characters elicited more positive nonverbal behavior (M = .64) than the black characters (M = -1.11), $t(54) = 8.39$, $p < .001$ (see rating procedure in Study 3, above). The pro-black condition contained clips in which white characters (M = -.94) elicited less positive nonverbal behavior than black characters (M = 1.11), t (53) = 8.66, p <.001. The control condition included clips in which white (M = -.20) and black characters (M = -.24) elicited nonverbal behavior that was roughly equivalent in positivity, $t(53) = .16$, p = .87.

Character ratings. Participants were asked to rate the extent to which they "liked" each of the characters they viewed following exposure to pro-White or pro-Black video clips. They were instructed to make these ratings on the basis of a "gut response." For each character, a neutral expression picture (downloaded from each show's website) was provided for participants, who made their ratings on a 1 (not at all liked) to 7 (extremely liked) scale. The ratings of the black characters were averaged as were the ratings of the white characters and the difference between these averages were indexed for each participant, representing the degree to which they favored the white over the black exemplars.

Self-reported prejudicial attitudes. The attitudes toward blacks scale (*S4*) is a reliable and established measure for assessing self-reported racial attitudes of white people toward black people. This scale consists of 10 negatively-phrased items (e.g., "I would rather not have black people live in the same apartment building I live in") and 10 positively-phrased items (e.g., "Black and white people are inherently equal"), each of which is to be rated on a 1 (strongly disagree) to 7 (strongly agree) scale. After reverse-coding and then averaging (i.e., dividing the sum by 20), higher scores on this scale indicate more negative attitudes toward black people.

Participants also completed the Asian modern racism scale (AMRS; *S5*)—this measure is not reported in the main text. It is a reliable and recently established scale for assessing anti-Asian attitudes. This scale includes 9 statements with which participants indicate their agreement (from -4, very strongly disagree to +4, very strongly agree). Examples of statements include "there are too many foreign students of Asian descent being allowed to attend college in America" and "it's good to live in a country where there are so many Asians." After reverse coding and then averaging, higher-scores on this measure are said to indicate greater anti-Asian sentiments.

Affective priming task. This task consisted of 96 trials in which a positive or negative target image followed a "subliminal" facial image (which was white, black, or Asian). Images of black, white, and Asian males and females were culled from the MacBrain stimulus set (*S6*), the Hess collection (*S7*), the JACFEE collection (*S8*), and our own collection. For each race, 4 male and 4 female photographs were selected, for a total of 24 facial images. Each facial image was paired with two positive and two negative target images.

Target images were selected on the basis of three requirements. Images selected were those that (a) were clearly positive or negative, (b) did not include violent or sexual material, and (c) did not include emotional facial expressions. The majority of the images were selected from the International Affective Picture System (IAPS; *S9*). The positive images selected from the IAPS included flowers, kittens, bunnies, dolphins, nature scenes and babies. The positive IAPS images were supplemented by public domain internet images depicting birds, desserts, and nature scenes. The negative images selected from the IAPS included cockroaches, spiders, snakes, car accidents, and a scar. The negative IAPS images were supplemented by public domain internet images depicting scorpions, demons, bats, and a skull and crossbones.

For each trial, participants were asked to focus their attention on a row of asterisks that would appear on the center of the monitor for 1 second. Subsequently, a facial image was chosen at random (without replacement) to appear for 12 milliseconds immediately prior to a masking image, which remained on the screen for 75 milliseconds. A target image replaced the masking image and remained on the screen until participants pressed one of two keys ("a" and "l", counterbalanced) to indicate that the target image was "good" or "bad"—their task was to respond as quickly and accurately as possible to the target image (for similar procedures see *S10, S11*). Reaction-time to each target image was recorded via DirectRT^{TM} software. For each participant, reaction-times were discarded if they were 2.5 standard deviations above their mean reaction time. Moreover, reaction times shorter than 100 milliseconds were discarded. Finally, analyses using log-transformed reaction-times reveal identical results to those obtained using the raw reaction-times; for that reason, all analyses reported in the main text and below regard raw reaction-times.

Following previous research (e.g., *S11, S12*), average response times to positive targets were subtracted from average response times to negative targets for each prime

category (e.g., "white"). See Table S5 for reaction-times to each prime-target combination.

Procedure

After completing informed consent, participants followed instructions that appeared on a computer monitor. As part of these instructions, participants were told that the experiment was about social concentration and, as such, they would be expected to ignore certain images.

Additional Statistics

There was a correlation between positivity of white associations and relative liking for white over black characters (see Table 2). And as compared to the control condition, exposure to pro-white nonverbal bias influenced both of these correlated variables. Consequently, it was appropriate to examine whether the effect of exposure to pro-white bias on relative liking was *mediated* by positive white associations. Such mediational analysis was not appropriate for the pro-black exposure condition because exposure to pro-black bias did *not* influence white associations (see Figure 2). Mediation analyses thus utilized a dummy-coded exposure variable that compared the pro-white exposure condition (1) to the control condition (0).

The dummy-coded exposure variable significantly and positively predicted white associations [$b = 51.35$, $t\,(36) = 2.62$, $p = .01$] and relative liking [$b = .80$, $t\,(36) = 2.94$, p .005]. When entered together as predictors, white associations positively predicted relative liking [this relationship was marginal; $b = .003$, $t\,(36) = 2.63$, p .11] and critically, the predictive power of the exposure variable dropped [to $b = .61$, $t\,(36) = 2.11$, p .04]. To formally assess the indirect effect of exposure to nonverbal bias on relative liking via white associations, a bias-corrected bootstrap mediation model was utilized, as is recommended for relatively small sample sizes (here the relevant n = 35; see Efron & Tibshirani, 1993; Preacher & Hayes, 2004; Shrout & Bolger, 2002). We used a directional test, which requires that the 5% cutoff value in the lower-tail of the bootstrap distribution of indirect effects be above zero to obtain significance. This bootstrap utilized 1,000 resamples of the original dataset, yielding 1,000 estimates of each path, including the indirect path. We found that this cutoff of the indirect effect of attributions to discrimination was, in fact, above zero (.01). Hence, the effect of exposure to pro-white nonverbal bias on relative liking for white over black characters is partially mediated by positive associations with white people.

We supplemented analyses on character liking with a 3 (exposure condition) x 2 (character race) mixed-model ANOVA with repeated measure on the last factor. A main effect of race revealed that black characters were generally rated higher than white characters, $F(1,53) = 8.35$, $p = .006$. However, this effect was qualified by a race by exposure condition interaction, $F(2,52) = 14.25$, $p < .001$. Follow-up t-tests revealed that white characters were rated higher by participants in the pro-white condition ($M = 4.54$) as compared to participants in the control condition ($M = 4.07$), $t(37) = 2.48$, $p = .02$, and

as compared to participants in the pro-black condition (M = 4.05), $t(33) = 1.99$, $p = .055$ (these latter two conditions did not differ: $t(36) = .11$, $p = .91$). Conversely, black characters were rated higher by participants in the pro-black condition (M = 5.13) as compared to participants in the control condition (M = 4.45), $t(36) = 3.38$, $p = .002$, and as compared to participants in the pro-white condition (M = 4.08), $t(33) = 4.08$, $p < .001$ (these latter two conditions did not differ: $t(37) = 1.50$, $p = .15$. In general, exposure to pro-white nonverbal bias lead to higher liking ratings of white characters (as compared to the other two conditions); conversely, exposure to pro-black nonverbal bias lead to higher liking ratings of black characters (as compared to the other conditions).

As described in the main text, exposure to pro-black nonverbal bias led to more positive self-reported attitudes toward blacks than did exposure to pro-white nonverbal bias or exposure to the control condition. Moreover, these effects did *not* extend to self-reported anti-Asian bias. Specifically, anti-Asian bias was *not* lower after exposure to pro-black nonverbal bias (M = 3.42) than after exposure to pro-white nonverbal bias (M = 3.49) nor as compared to the control condition (M = 3.62), $F(2,53) = 1.96$, $p = .75$.

References

S1. N. Ambady, R. Rosenthal, *Psychol. Bull.,* **111***,* 256 (1992).

S2. A. G. Greenwald, D. E. McGhee, J. L. K. Schwartz, *J. Pers. Soc. Psychol.* **74**, 1464 (1998).

S3. A. G. Greenwald, B. A. Nosek, M. R. Banaji, *J. Pers. Soc. Psychol.* **85**, 197 (2003).

S4. J. C. Brigham, *J. Appl. Soc. Psychol.*, **23**, 1933 (1993).

S5. L. S. S. Hing, W. Li, M. P. Zanna, *J. Exp. Soc. Psychol.,* **38**, 71 (2002).

S6. N. Tottenham, A. Borscheid, K. Ellertsen, D. J. Marcus, C. A. Nelson, paper presented at Cognitive Neuroscience Society Annual Meeting, San Francisco, CA, 15 April 2002 (2002).

S7. M. Beaupre, U. Hess, *J. Cross Cult. Psychol.,* **36,** 355 (2005).

S8. D. Matsumoto, P. Ekman, "Japanese and Caucasian Facial Expressions of Emotion (JACFEE)" (San Francisco: Author, 1988).

S9. P. J. Lang, M. M. Bradley, B. N. Cuthbert, "International affective picture system (IAPS): Instruction manual and affective ratings" (Gainesville, FL: Author, 1999).

S10. R. H. Fazio, J. R. Jackson, B. C. Dunton, C. J. Williams, *J. Pers. Soc. Psychol.* **69**, 1013 (1995).

S11. S. Sinclair, B. S. Lowery, C. D. Hardin, A. Colangelo, *J. Pers. Soc. Psychol*., **89**, 583 (2005).

S12. B. S. Lowery, C. D. Hardin, S. Sinclair, *J. Pers. Soc. Psychol.,* **81,** 842 (2001).

Table S1. *Character ratings for Study 1.* Nonverbal and verbal favoritism reflect ratings on -3 to +3 scales. Attractiveness, sociability, kindness, and intelligence ratings reflect ratings from 1 to 7 scales.

Character	Show	Race	Nonverbal Favoritism	Verbal Favoritism	Attractive	Sociable	Kind	Intelligent
"Bones" Brennan	*Bones*	White	-.098	-.02	5.69	4.75	4.75	5.88
Camille Saroyan	*Bones*	Black	.067	-.33	5.69	5.31	5.19	5.56
Gil Grissom	*CSI*	White	.227	.05	3.81	3.44	3.50	3.50
Warrick Brown	*CSI*	Black	.055	.09	5.25	4.75	4.63	4.63
Calleigh Duquesne	*CSI: Miami*	White	.129	.09	5.81	4.81	3.94	3.94
Alexx Woods	*CSI: Miami*	Black	.028	-.10	4.81	4.38	4.19	4.19
Jason Street	*Friday Night Lights*	White	.632	.41	5.69	4.94	4.50	4.19
"Smash" Williams	*Friday Night Lights*	Black	-.220	-.73	5.81	5.94	4.94	4.44
Casey	*Greek*	White	.239	.23	4.00	5.31	5.75	5.13
Ashleigh	*Greek*	Black	.057	-.13	6.06	5.56	4.81	4.69
Rusty	*Greek*	White	.418	.03	4.89	4.81	4.51	4.90
Calvin	*Greek*	Black	.239	.74	5.19	5.81	5.13	5.00
Derek Shepherd	*Grey's Anatomy*	White	.448	.53	6.00	5.69	5.13	6.19
Richard Webber	*Grey's Anatomy*	Black	-.142	.17	3.49	4.25	4.94	6.25
Mark Sloan	*Grey's Anatomy*	White	-.033	-.09	5.63	4.25	4.25	5.94
Preston Burke	*Grey's Anatomy*	Black	.285	.51	4.19	4.31	4.38	6.19
Addison Montgomery	*Grey's Anatomy*	White	.388	.25	5.94	4.81	5.44	5.69
Miranda Bailey	*Grey's Anatomy*	Black	-.009	.00	5.44	4.06	3.94	6.06
Claire	*Heroes*	White	-.015	.49	6.06	4.69	4.75	4.75
Simone	*Heroes*	Black	-.467	-.06	5.06	3.56	4.06	4.88
Gregory House	*House*	White	-.158	-.02	5.44	4.19	3.94	5.63
Eric Foreman	*House*	Black	-.273	-.05	5.25	4.81	4.69	6.13
James Garcia	*Reno 911!*	White	-.401	.15	2.63	4.06	3.25	2.81
S. Jones	*Reno 911!*	Black	-.338	.23	3.50	5.94	4.88	4.13
Trudy Wiegal	*Reno 911!*	White	-.034	.32	2.88	4.69	4.13	3.13
Raineesha Williams	*Reno 911!*	Black	-.162	-.06	2.88	6.06	4.81	3.25
Rob	*Rob and Big*	White	.235	-.01	3.88	5.44	4.38	3.50
Big	*Rob and Big*	Black	.250	.15	2.81	6.06	4.88	3.81
J. D.	*Scrubs*	White	.480	.24	4.94	6.00	5.88	5.50
Turk	*Scrubs*	Black	-.106	.20	5.25	6.25	5.81	5.56

Table S2. *Study 1 correlations among character ratings. p*-values in italics. N = 30

	Nonverbal Favoritism	Verbal Favoritism	Attractiveness	Sociability	Kindness
Nonverbal favoritism					
Verbal favoritism	.47 *.01*				
Attractiveness	.21 *.27*	.21 *.27*			
Sociability	.23 *.23*	.04 *.84*	-.02 *.93*		
Kindness	.32 *.08*	.15 *.43*	.24 *.21*	.74 *<.001*	
Intelligence	.16 *.40*	.04 *.85*	.62 *<.001*	-.25 *.18*	.24 *.20*

Table S3. *Calculation of IAT scores.* Blocks w-p and b-p are each subdivided into an early ("practice") block and a late block. Average reaction time in the early portion of the w-p block (*wpe*) is subtracted from average reaction time in the early portion of the b-p block (*bpe*) and this difference is divided by the standard deviation of early trials (SD_e). Similarly, average reaction time in the late portion of the w-p block (*wpl*) is subtracted from average reaction time in the late portion of the b-p block (*bpl*) and this difference is divided by the standard deviation of late trials (SD_l). These two standardized difference scores are then averaged for the final IAT score. The by-study calculations are illustrated with grand means and overall standard deviations.

	Equation	Average RT w-p	Average RT b-p
General Equation	$[(bpe - wpe)/ SD_e + (bpl - wpl)/ SD_l]/2$	(*wpe*, *wpl*)	(*bpe*, *bpl*)
Study 2	[(1112.21 – 795.89)/478.55 + (875.42 – 696.41)/369.75]/2	746.15	993.81
Study 3a	[(1133.35 – 806.01)/506.76 + (855.19 – 705.50)/339.98]/2	755.76	984.27
Study 3b	[(1110.19 – 768.44)/500.91 + (845.84 – 684.04)/353.69]/2	726.24	978.02

Table S4. *Study 3b characters by condition*.

Character Race	Character Role	Pro-black exposure	Pro-white exposure	Control Condition
Black				
	Doctor	Bailey *Grey's Anatomy*	Same	Same
	Doctor	Burke *Grey's Anatomy*	Turk *Scrubs*	Both
	Doctor	Webber *Grey's Anatomy*	Same	Same
	Doctor	Foreman *House, MD*	Same	Same
	Detective	David *The Closer*	Same	Same
	Fraternity Member	Calvin *Greek*	Same	Same
	Reality Roommate	Tyrie *The Real World*	Same	Same
	Reality Roommate	Stephen *The Real World*	Same	Same
White				
	Doctor	House *House, MD*	Same	Same
	Doctor	Jack *Lost*	J.D. *Scrubs*	Both
	Doctor	Allison *House, MD*	Same	Same
	Investigator	Natalia *CSI: Miami*	Yelina *CSI: Miami*	Neither
	Castaway	Claire *Lost*	Same	Same
	Superhero	Matt *Heroes*	Claire *Heroes*	Both
	Sick Supermodel	Alex (guest) *House, MD*	Same	Same
	Investigator/Doctor	Ryan *CSI: Miami*	Meredith *Grey's Anatomy*	Neither

Table S5. *Average reaction times in Study 4.* Standard deviations in italics.

	Positive target images	Negative target images
White faces	676.83 *133.50*	669.76 *137.91*
Black faces	671.60 *131.23*	666.90 *133.44*
Asian faces	669.35 *128.86*	666.62 *129.36*