

On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals

Nilanjana Dasgupta
New School University

Anthony G. Greenwald
University of Washington

Two experiments examined whether exposure to pictures of admired and disliked exemplars can reduce automatic preference for White over Black Americans and younger over older people. In Experiment 1, participants were exposed to either admired Black and disliked White individuals, disliked Black and admired White individuals, or nonracial exemplars. Immediately after exemplar exposure and 24 hr later, they completed an Implicit Association Test that assessed automatic racial attitudes and 2 explicit attitude measures. Results revealed that exposure to admired Black and disliked White exemplars significantly weakened automatic pro-White attitudes for 24 hr beyond the treatment but did not affect explicit racial attitudes. Experiment 2 provided a replication using automatic age-related attitudes. Together, these studies provide a strategy that attempts to change the social context and, through it, to reduce automatic prejudice and preference.

Prejudice reduction has been a hot topic in social psychology for several decades. Despite the abundant research devoted to this topic, reviews of the literature reveal that attempts to reduce prejudice and discrimination have, at best, yielded mixed findings (Duckitt, 1992; Harding, Kutner, Proshansky, & Chein, 1954; Monteith, Zuwerink, & Devine, 1994; Stephan, 1985). Over the past few decades, social psychologists have developed a number of interventions aimed at alleviating prejudice. Classic interventions include the promotion of interpersonal contact among members of high- and low-status groups (Amir, 1969; Cook, 1985; Brewer & Miller, 1988) and the promotion of cognitive consistency between general egalitarian values and attitudes toward specific groups (Gaertner & Dovidio, 1986; Katz & Hass, 1988; Katz, Wackenhut, & Hass, 1986; Roakeach, 1973). Contemporary approaches focus on the motivated suppression of negative stereotypes (Macrae, Bodenhausen, Milne, & Jetten, 1994; Macrae, Bodenhausen,

Milne, & Wheeler, 1996) or replacement of automatic responses based on culturally shared stereotypes with effortful responses based on personal beliefs (Devine, 1989; Monteith, 1993; Monteith, Devine, & Zuwerink, 1993; Monteith et al., 1994).

The reduction of prejudice has typically been viewed as an effortful and self-conscious relearning process. As such, most of the research described above has focused on changing attitudes within people's awareness and volitional control. Few have attempted to modify implicit and automatic forms of prejudice that are less available to introspection and control. In fact, the activation of automatic beliefs and attitudes has been described as an inescapable habit that occurs despite attempts to bypass or ignore it (Bargh, 1999; Devine, 1989). Current models of prejudice and stereotype reduction argue that prejudice-free responses require perceivers to be aware of their bias; to be motivated to change their responses because of personal values, feelings of guilt, compunction, or self-insight (Allport, 1954; Devine, Monteith, Zuwerink, & Elliot, 1991; Myrdal, 1944); and to possess cognitive resources necessary to develop and practice correction strategies (Blair & Banaji, 1996; Fiske & Neuberg, 1987; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). We propose that effortful processes may not be the only way to moderate implicit prejudice. In support of this idea, several new theories of attitude and empirical demonstrations suggest that attitudes are multifaceted evaluations, shaped by a number of factors only one of which is explicit motivation (Jarvis, 1998; Petty & Jarvis, 1998; Smith & Zárate, 1992; Wilson, Lindsey, & Schooler, 2000).

Evidence for the Malleability of Attitudes

Attitudes that are expressed at any given time appear to depend on various factors such as people's cognitive resources, current focus of attention, motivation, goals, and contextual cues (Petty & Jarvis, 1998; Smith & Zárate, 1992; Wilson et al., 2000). Specifically, empirical evidence has demonstrated that perceivers' beliefs and attitudes vary as a function of the following factors: (a) cues

Nilanjana Dasgupta, Department of Psychology, New School University; Anthony G. Greenwald, Department of Psychology, University of Washington.

This research was supported by National Science Foundation Grants SBR-9422242 and SBR-9710172 and by National Institute of Mental Health Grants MH-41328, MH-01533, and MH-57672. These data were presented at the meeting of the American Psychological Society, Denver, Colorado, June 1999. We thank Mahzarin Banaji, Shelly Farnham, Elizabeth Haines, Debbie McGhee, Deborah Mellott, Brian Nosek, Jaihyun Park, Jane Swanson, and Eric Uhlmann for their valuable feedback. We are also grateful to Richard Abrams, Shelly Farnham, Jennifer Guinn, Debbie McGhee, Deborah Mellott, Jane Swanson, and Eric Uhlmann for their assistance with data collection.

Correspondence concerning this article should be addressed to Nilanjana Dasgupta, Department of Psychology, New School University, 65 Fifth Avenue, New York, New York 10003 or to Anthony G. Greenwald, Department of Psychology, University of Washington, Seattle, Washington 98195-1525. Electronic mail may be sent to dasguptn@newschool.edu or to agg@u.washington.edu.

in the social context (e.g., Feldman & Lynch, 1988; McGuire & Padawer-Singer, 1976; Schwarz & Bless, 1992; Smith, Fazio, & Cejka, 1996; for a review see Wilson & Hodges, 1992); (b) the perceivers' current mood (Forgas, 1992; Petty, Schumann, Richman, & Strathman, 1993; Schwarz & Clore, 1983); (c) their present thoughts and feelings (Chaiken & Yates, 1985; Wilson & Hodges, 1992); and (d) prior expectancies (Cialdini, 1976; McFarland, Ross, & Conway, 1984; Tesser & Danheiser, 1978).

Furthermore, empirical evidence for the malleability of implicit beliefs and attitudes has also been accumulating. Studies on construct accessibility have shown that a construct rendered temporarily accessible in one situation (by exposure to traits or behaviors) can shape people's attitude- and belief-based responses expressed in a subsequent situation without their awareness (Banaji, Hardin, & Rothman, 1993; Bargh & Pietromonaco, 1982; Devine, 1989; Herr, 1986; Higgins, Rholes, & Jones, 1977; Lewicki, 1985; Sherman, Mackie, & Driscoll, 1990; Srull & Wyer, 1979; Strack, Schwarz, Bless, & Kuebler, 1993). The judgments measured in these studies were implicit in the sense that perceivers were unaware that their responses were shaped by previously encountered information. However, the specific dependent measures all involved self-reports that were clearly under participants' voluntary control.

A few recent studies have suggested that even attitudes and beliefs that are difficult to bring under volitional control (i.e., that are "implicit" in a different way) may be sensitive to (a) perceivers' task goals and expectancies (Blair & Banaji, 1996; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997); (b) practice or training (Kawakami et al., 2000); (c) automatic motives (Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Moskowitz, Salomon, & Taylor, 2000); and (d) preexisting individual differences in attitudes, beliefs, and motivation to respond without prejudice (Amodio, Harmon-Jones, & Devine, 2000; Lepore & Brown, 1997). The first two sets of studies cited above demonstrated the plasticity of automatic beliefs and attitudes by teaching participants specific strategies to combat automatic responses and by relying on their motivation and effort to implement them.

Goals of the Present Research

We sought to test whether automatic negative attitudes can be temporarily modified using a different technique that focuses on changing the social context that people inhabit rather than by directly manipulating their goals and motivations. Specifically, if perceivers are immersed in situations that provide frequent exposure to admirable members of stigmatized groups (e.g., famous African Americans) and disliked members of valued groups (e.g., infamous European Americans), their automatic intergroup attitudes may shift in important ways. For such an intervention to be useful, it must not only produce an immediate effect on intergroup attitudes but must also demonstrate longer lasting effects. For it to be efficient, the strategy must not be limited to one social group but must be generalizable to other groups targeted by prejudice.

In Experiment 1, we developed an intervention that tested whether repeated exposure to images of famous and admired African Americans as well as infamous and disliked European Americans can moderate automatic racial attitudes. Moreover, we assessed whether the predicted shift in automatic race bias was stable enough to endure for a substantial period of time after the

manipulation. Experiment 2 examined the generalizability of this strategy by testing its influence on automatic attitudes toward younger and older people.

Measuring Automatic Attitudes: A Brief Review of the Implicit Association Test

In all experiments reported here, implicit attitudes were measured using the Implicit Association Test (IAT), a task in which speed of response is used to assess the relative strength with which attitude objects are associated with positive versus negative evaluations. In this task, developed by Greenwald, McGhee, and Schwartz (1998), participants classify stimuli representing two target concepts (e.g., flowers and insects, or Black and White groups) and evaluative attributes (e.g., good and bad words) using two designated keys. When the IAT is used to measure racial attitudes, people typically respond more quickly and easily if pleasant attributes share the same response key with White racial stimuli and unpleasant attributes share the same key with Black racial stimuli than vice versa (Dasgupta, McGhee, Greenwald, & Banaji, 2000; Greenwald et al., 1998; Ottaway, Hayden, & Oakes, 2001).

At its core, the IAT is similar to an older and more commonly used indirect measure—the evaluative priming task (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Both tasks are based on the assumption that if an attitude object evokes a particular evaluation (positive or negative), it will facilitate responses to other evaluatively congruent and co-occurring stimuli. For both tasks, response facilitation is interpreted as a measure of the strength of association between the object and evaluation. Evaluative priming measures, which have been circulating in the literature for the past 15 years, have been accepted by many social psychologists as revealing automatic attitudes (Bargh, Chaiken, Gendler, & Pratto, 1992; Chen & Bargh, 1999; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1986; Fazio, Jackson, Dunton, & Williams, 1995; Glaser & Banaji, 1999). By comparison, the IAT is a newer task, but given its conceptual similarity to evaluative priming and the growing evidence supporting its validity, we argue that the IAT should also be considered a measure of automatic attitude. To be sure, not everybody agrees that evaluative associations captured by IATs or priming tasks, for that matter, represent attitudes, as is clear from a few of the articles in this special issue of the *Journal of Personality and Social Psychology*. However, we believe that a sufficient number of articles (both published and in press) now provide reasonably strong empirical evidence for the construct validity of the IAT as an automatic attitude measure. These findings are summarized below.

Construct validity of the IAT has been assessed using various tests of known group differences. These have shown that implicit group preferences captured by the IAT reliably predict people's membership in various groups on the basis of the following characteristics: (a) race/ethnicity (Dasgupta et al., 2000; Greenwald et al., 1998; Ottaway et al., 2001); (b) nationality (Greenwald et al., 1998; Rudman, Greenwald, Mellott, & Schwartz, 1999); (c) age (Mellott & Greenwald, 1999); (d) sex (Farnham & Greenwald, 1999); (e) sexual orientation (Dasgupta, 2001); (f) academic preferences (e.g., entomology majors and science majors; Citrin & Greenwald, 1998; Nosek, Banaji, & Greenwald, 1998); and (g)

dietary preferences (e.g., vegetarian and nonvegetarian; Swanson, Rudman, & Greenwald, 2001).

The interpretation that IAT effects reveal automatic attitudes rests critically on tests that rule out threats to internal validity, the most common being that automatic group preference captured by this task may stem from greater familiarity with stimuli representing one group over the other. This potential confound was addressed by Dasgupta et al. (2000) and Ottaway et al. (2001) specifically with regard to racial attitudes. In their studies, even when Black and White stimuli were selected to be equally familiar and when familiarity with racial stimuli was statistically controlled, participants continued to exhibit implicit White preference (also see Rudman, Greenwald, et al., 1999, for another way of ruling out this potential confound).

Tests of convergent validity have revealed substantial correlations between IATs and automatic priming tasks—after correcting for measurement error—suggesting that both procedures capture the strength of automatic associations (Cunningham, Preacher, & Banaji, 2001; Rudman & Kilianski, 2000). Moreover, IAT-assessed implicit prejudice exhibited by White participants has been shown to correlate significantly with activation of the amygdala, a subcortical structure associated with emotional learning and evaluation ($r = .58$) and eye-blink startle responses ($r = .56$) when participants were shown Black compared with White faces. At the same time, amygdala activity, assessed by functional magnetic resonance imaging (fMRI), and eye-blink measures were not correlated at statistically significant levels with explicit measures of racial attitude and belief, including the Modern Racism Scale (Phelps et al., 2000).

Finally, new data have provided support for the predictive validity of the IAT. McConnell and Leibold (2001) found that implicit racial attitudes as measured by the IAT predicted participants' nonverbal behavior toward Black and White experimenters (both as assessed by trained judges reviewing the videotaped interactions and by the experimenters themselves). Specifically, implicit pro-Black attitudes were associated with more positive interactions with the Black compared with the White experimenter. Explicit attitudes (as assessed by feeling thermometers and semantic differentials) were more modestly correlated with interracial behavior. In other research, Sargent and Theil (2000) found that implicit racial attitudes as measured by the IAT predicted participants' choice of work partner (Black or no race specified)—but only under conditions of high attributional ambiguity.

Experiment 1

In this experiment, we first sought to test whether repeated exposure to admired African American and disliked European American exemplars would produce any decrement in the magnitude of automatic preference for European over African Americans. Second, we sought to determine whether the reduced race preference effect, if obtained, would endure for at least 24 hr. Third, we assessed whether the exemplar strategy would affect explicit or self-reported attitudes. We expected that self-reported attitudes would remain unchanged despite exemplar exposure because participants had time to reflect on the typicality of the famous and infamous exemplars and possessed the mental resources to correct their explicit responses if they so desired (Allport, 1954; Martin, 1986; Weber & Crocker, 1983).

Participants first completed a task that was ostensibly a measure of "general knowledge." In this task, they were shown pictures of either admired Black and disliked White individuals (pro-Black exemplar condition), disliked Black and admired White individuals (pro-White exemplar condition), or nonracial exemplars (control condition). Participants' task was to correctly identify the person (or object) seen in the pictures. After exemplar exposure, implicit racial attitudes were measured with the IAT and explicit racial attitudes were assessed with feeling thermometers and semantic differential scales. We also examined whether the automatic attitude responses produced by these exemplars would endure for 24 hr by reassessing people's attitudes the next day.

Method

Participants

Students ($N = 48$; 25 women, 23 men) enrolled in introductory psychology courses at the University of Washington participated in exchange for extra course credit. There were 31 Caucasian and 17 Asian participants. Before signing up for the experiment, participants were informed that the study required two 0.5-hr sessions separated by 24 hr. All students who participated in the first session returned for the second session.

Materials

Selection of exemplars. Pictures of 40 well-known Black and White individuals were culled from the Internet using various search engines. Of the 40 pictures, 10 were in each of the following categories: (a) admired Black individuals (e.g., Denzel Washington), (b) admired White individuals (e.g., Tom Hanks), (c) disliked Black individuals (e.g., Mike Tyson), and (d) disliked White individuals (e.g., Jeffrey Dahmer).¹ In addition, pictures of 10 flowers (e.g., sunflower) and 10 insects (e.g., mosquito) were also gathered for the control condition. All pictures were converted into a standard format of 104×202 pixels in dimension and 256-color gray scale.

Automatic attitude measure. Automatic attitudes were measured with the IAT. In this task, racial groups were represented by five African American and five European American first names (e.g., Jamal, Justin) previously used by Dasgupta et al. (2000) and Greenwald et al. (1998). The evaluative attribute was represented by five pleasant and five unpleasant words (e.g., *paradise*, *poison*) selected from Bellezza, Greenwald, and Banaji (1986). Both the IAT and the exemplar exposure task were administered on PC-type (80486 processor) desktop computers equipped with color monitors and Windows 95 operating systems.² All names and evaluative words are listed in Appendix A.

Feeling thermometers. Two feeling thermometers assessed the favorability of participants' explicit feelings about White and Black Americans. They were asked to mark an appropriate position on a picture of a thermometer numerically labeled at 10° intervals from 0° (*cold* or *unfavorable*) to 99° (*warm* or *favorable*) to indicate their attitudes.

Semantic differentials. Participants also completed five semantic differential scales for each group. These 7-point scales (-3 to $+3$) were anchored at either end by polar-opposite adjective pairs: ugly–beautiful, bad–good, unpleasant–pleasant, dishonest–honest, and awful–nice.

¹ All admired and disliked exemplars in this experiment were men. Female exemplars were not used in this study because we were unable to generate a sufficient number of names of well-known yet strongly disliked Black women.

² This computerized experiment was administered using Inquisit, an experimental program developed by Sean C. Draine (1998).

Procedure

Exposure to exemplars. The exemplar task was presented as a general knowledge test assessing participants' familiarity with famous and infamous Americans (in the pro-Black and pro-White conditions) or with flowers and insects (in the control condition). In the first two conditions ($n = 18$ and 15 , respectively), participants viewed pictures of Black and White males, twice each, in the first block of 40 trials. Each picture was accompanied by the person's name together with a correct and incorrect description of him. For instance, Martin Luther King's picture was accompanied by a correct description that read "Leader of the Black Civil Rights movement in the 1960s" and an incorrect description that read "Former Vice President of the United States." Correct and incorrect descriptions were always matched in valence so that in the event of an error, the stimulus individual would be remembered in the desired positive or negative fashion. The picture and name were centered on the screen and the two descriptions appeared side by side below the picture. See Appendix B for a complete list of exemplar names and descriptions.

In the control condition ($n = 15$) participants saw pictures of flowers and insects twice each for the first block of 40 trials. Each picture was accompanied by a true and false name of the stimulus that appeared side by side below the picture. See Appendix C for a complete list of flowers and insects used in this experiment.

In all three conditions, the participants' task was to identify the correct description by pressing the *A* key to select the description on the left or the *5* key on the numeric keypad to select the description on the right. Half of the correct descriptions appeared on each side. Incorrect identifications were followed by error feedback (i.e., the word *error*).

In the second block of 40 trials, participants were shown exemplar names (without any pictures) together with the instruction to classify them as Black or White (in the experimental conditions) or as flowers or insects (in the control condition). This was done to ensure that participants in the experimental conditions could accurately recall the race of the exemplars. They were again instructed to use the *A* and *5* keys for their classifications. For half of the participants, the *A* key was used to classify Black names (or flowers) and the *5* key to classify White names (or insects). For the remaining participants, key assignment was reversed.

IAT. After the exemplar task, participants completed the IAT, during which they were instructed to categorize four types of stimuli (Black and White names, pleasant and unpleasant words) using two designated response keys on a computer keyboard. When highly associated targets and attributes share the same response key, participants tend to classify them quickly and easily, whereas when weakly associated targets and attributes share the same response key, participants tend to classify them more slowly and with greater difficulty. Given the subtle but pervasive presence of White preference and anti-Black sentiments in mainstream American culture, we anticipated faster IAT performance when White and pleasant stimuli shared the same key and when Black and unpleasant stimuli shared the other key (abbreviated as White + pleasant and Black + unpleasant, respectively). By contrast, we expected substantially slower performance for opposite combinations of stimuli (Black + pleasant and White + unpleasant). Automatic White preference was measured as the difference in mean response latencies for pro-White stimulus combinations (White + pleasant and Black + unpleasant) versus pro-Black stimulus combinations (Black + pleasant and White + unpleasant).³

Explicit attitude measures. Next, participants filled out feeling thermometers and semantic differential scales. They were assured in both oral and written instructions that their responses would remain completely anonymous.⁴ Finally, participants were told that the experimental session was over and were asked to return to the laboratory 24 hr later. On the 2nd day, they completed the IAT and self-report measures again without being reminded of the exemplars seen the previous day.

Design

The experimental design was a 3 (type of exemplars: pro-Black, pro-White, control) \times 2 (IAT combinations: Black + pleasant vs. White + pleasant) \times 2 (time delay: immediate vs. delayed administration of dependent measures) mixed factorial. The first factor was varied between subjects and the remaining two varied within subjects. In addition, two other between-subjects counterbalancing factors were included: (a) order of IAT combinations (Black + pleasant first vs. White + pleasant first) and (b) key assignment for exemplar classification (Black with left key vs. White with left key or flower with left key vs. insect with left key).

Results and Discussion

Data Preparation

Data collection blocks of the IAT were retained and practice blocks were discarded. Additionally, the first trial from each data collection block was deleted because response latencies were typically longer. To correct for anticipatory responses and momentary inattention, latencies less than 300 ms and greater than 3000 ms were recoded as 300 and 3000 ms, respectively. These latencies were then log transformed to normalize the distribution (see Dasgupta et al., 2000; and Greenwald et al., 1998, for a similar procedure). Figure 1 illustrates the influence of pro-Black exemplars on automatic racial attitudes, both immediately after exemplar exposure (Panel A) and after a 24-hr delay (Panel B).

IAT Administered Immediately After Exemplar Exposure

As shown in Figure 1, Panel A, results revealed that exposure to pro-Black exemplars had a substantial effect on automatic racial associations (or the IAT effect).⁵ The magnitude of the automatic White preference effect was significantly smaller immediately after exposure to pro-Black exemplars (IAT effect = 78 ms; $d = 0.58$) compared with nonracial exemplars (IAT effect = 174 ms; $d = 1.15$), $F(1, 31) = 6.79$, $p = .01$; or pro-White exemplars (IAT effect = 176 ms; $d = 1.29$), $F(1, 31) = 5.23$, $p = .029$. IAT effects in control and pro-White conditions were statistically comparable ($F < 1$).⁶

IAT Administered 24 Hours After Exemplar Exposure

Panel B of Figure 1 illustrates the response latency data 24 hr after exemplar exposure. Compared with the control condition, the

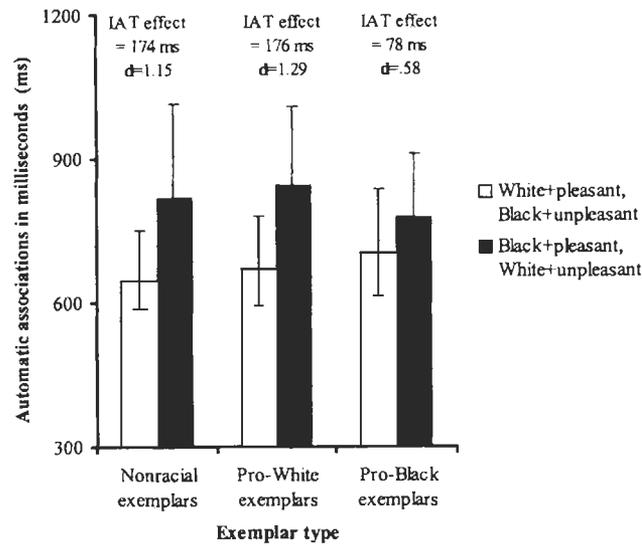
³ Detailed descriptions of the IAT are available in Greenwald et al. (1998).

⁴ Because the primary prediction in this experiment focused on the malleability of automatic attitudes, the IAT was administered first followed by the explicit measures.

⁵ IAT effects can be computed by subtracting the mean latency for White + pleasant and Black + unpleasant blocks from the mean latency for Black + pleasant and White + unpleasant blocks. Thus, positive difference scores indicate that pleasant attributes are more strongly associated with White than Black and unpleasant attributes with Black than White.

⁶ These effect sizes were computed using the pooled standard deviations of pro-Black and pro-White blocks as the denominator and the difference between the two blocks as the numerator.

A



B

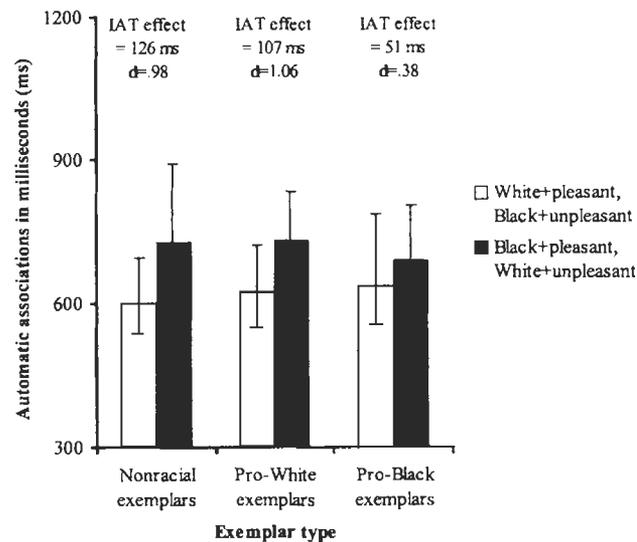


Figure 1. A: Effect of pro-Black versus pro-White exemplars on implicit race associations immediately after exemplar exposure. Error bars represent the standard deviation of responses in each condition. B: Effect of pro-Black versus pro-White exemplars on implicit race associations 24 hr after exemplar exposure. IAT = Implicit Association Test.

magnitude of the IAT effect in the pro-Black condition remained significantly diminished 1 day after encountering admired Black and disliked White images (IAT effects = 126 ms vs. 51 ms, respectively; d s = 0.98 vs. 0.38, respectively), $F(1, 31) = 4.16$, $p = .05$. Similarly, compared with the pro-White condition, the IAT effect in the pro-Black exemplar condition remained substantially smaller as well (IAT effects = 107 vs. 51 ms, respectively; d s = 1.06 vs. 0.38, respectively), $F(1, 31) = 3.67$, $p = .065$.

Explicit Measures Administered Immediately After Exemplar Exposure

Feeling thermometers yielded an evaluative rating for each group in which higher scores represented more favorable attitudes. Semantic differential scales were scored by averaging the five items to create an attitude index for each racial group in which higher scores indicated more favorable attitudes. Participants ex-

pressed greater liking for White compared with Black Americans on feeling thermometers ($M_s = 71^\circ$ and 66° ; $d = 0.24$), $F(1, 45) = 6.35$, $p = .015$; and semantic differential scales ($M_s = 0.58$ and 0.45 ; $d = 0.15$), $F(1, 45) = 4.05$, $p = .05$. Exposure to pro-Black exemplars did not reduce self-reports of White preference on either measure. For each exemplar condition, the effect size difference between Black and White ratings on the feeling thermometer was as follows: $d = 0.31$ (control); $d = 0.03$ (pro-White); $d = 0.39$ (pro-Black); Exemplar Type \times Racial Group interaction: $F(1, 46) = 1.91$, $p = .17$. The difference between Black and White ratings on the semantic differential index was $d = 0.17$ (control); $d = 0.12$ (pro-White); $d = 0.13$ (pro-Black); Exemplar Type \times Racial Group interaction: $F(1, 46) = 0.06$, $p = .81$.

Explicit Measures Administered 24 Hours After Exemplar Exposure

The pattern of explicit data remained stable 24 hr after encountering the exemplars. Participants continued to report slightly more favorable attitudes toward White compared with Black Americans overall, both on feeling thermometers ($M_s = 68^\circ$ and 65° respectively; $d = 0.18$), $F(1, 42) = 5.53$, $p = .02$; and semantic differential scales ($M_s = .64$ and $.48$; $d = .19$), $F(1, 42) = 7.39$, $p = .009$. As before, exposure to pro-Black exemplars did not reduce race preference. Specifically, the differences between Black and White ratings on feeling thermometers were as follows: $d = 0.34$ (control); $d = 0.05$ (pro-White); $d = 0.27$ (pro-Black); Exemplar Type \times Racial Group interaction: $F(1, 43) = 0.57$, $p = .45$. For semantic differential scales, effect sizes of differential race ratings were $d = 0.21$ (control); $d = 0.11$ (pro-White); $d = 0.25$ (pro-Black); Exemplar Type \times Racial Group interaction: $F(1, 43) = 1.10$, $p = .30$.

Relationship Between Implicit and Explicit Measures

Correlations between explicit and implicit attitude measures revealed nonsignificant relationships between automatic attitudes assessed by the IAT and explicit attitudes assessed by feeling thermometers and semantic differential scales ($r = .19$, $p = .21$, and $r = .12$, $p = .45$, respectively). The two explicit measures were significantly correlated with one another, both for ratings for Black Americans ($r = .43$, $p = .004$) and White Americans ($r = .51$, $p < .0009$).

Test-Retest Reliability

Correlations revealed that participants' responses on all measures remained remarkably stable across the two experimental sessions: (a) IATs: $r = .65$, $p < .0009$; (b) feeling thermometers: $r = .88$, $p < .009$ (White ratings), $r = .87$, $p < .0009$ (Black ratings); (c) semantic differentials: $r = .80$, $p < .0009$ (White ratings), $r = .78$, $p < .0009$ (Black ratings).

In sum, Experiment 1 showed that the magnitude of automatic White preference was significantly smaller when pro-Black images rather than nonracial or pro-White images were made salient. This effect was not only evident immediately after exemplar exposure, but also endured for 24 hr. Moreover, although pro-Black images reduced automatic White preference, pro-White

images did not exacerbate the effect—the magnitude of race preference remained quite similar regardless of whether participants were reminded of pro-White or nonracial exemplars. We speculate that perhaps pro-White exemplars had been chronically accessible to perceivers even in the control condition; thus additional exposure to the same type of images produced no further increase in automatic White preference.

By comparison, explicit attitude responses revealed a different pattern of data: They consistently revealed slight preference for European Americans over African Americans regardless of the types of images participants had seen previously. These two distinct patterns of data, together with our correlational findings showing small and nonsignificant correlations between explicit and implicit measures, support the argument that implicit and explicit attitudes are at least partially dissociated (Devine & Monteith, 1999; Dovidio et al., 1997; Fazio et al., 1995; Greenwald & Banaji, 1995; Monteith et al., 1993; Wilson et al., 2000). Moreover, they suggest that different underlying processes may be responsible for changes in implicit versus explicit attitudes (see the General Discussion for more details).

Experiment 2

A number of studies have shown that social perceivers tend to express implicit positive attitudes toward young people and relatively negative attitudes toward older people (Butler, 1980; Mellott & Greenwald, 1998, 1999; Perdue & Gurtman, 1990). As in the case of race bias, implicit forms of age bias have been documented even when explicit attitudes toward older people are fairly egalitarian. Experiment 2 attempted to test whether automatic youth preference can be diminished using the exemplar strategy developed in the previous experiment. To that end, participants first completed an ostensible general knowledge task in which they were either exposed to pro-elderly or pro-young images, after which implicit and explicit age-related attitudes were measured using the IAT, feeling thermometers, and semantic differential scales.

Method

Participants

Students ($N = 26$; 21 women, 5 men) enrolled in introductory psychology courses at the University of Washington participated for extra course credit. There were 14 Asian, 10 Caucasian, and 2 African American participants.

Materials and Procedure

Selection of exemplars. Pictures of 40 famous old and young individuals were culled from the Internet, of which 10 belonged to each of the following four categories: (a) admired old individuals (e.g., Mother Teresa, Albert Einstein), (b) admired young individuals (e.g., Sarah McLachlan, Ben Affleck), (c) disliked old individuals (e.g., Tammy Faye Bakker, Bob Packwood), and (d) disliked young individuals (e.g., Tonya Harding, Andrew Cunanan). Within each category, there were 4 women and 6 men (see Appendix D for names and descriptions).

Half the participants viewed images of admired young and disliked old people (pro-young condition), and the remaining viewed images of admired old and disliked young people (pro-elderly condition). We did not

have an additional control group in this study. Pictures and accompanying descriptions were displayed using the same procedure as in Experiment 1.

IAT. Participants completed an IAT in which 20 first names were used to represent old and young people, with an equal number of male and female names in each age category. Old people were represented by 10 names (e.g., Agatha, Albert) obtained from Civil War–related Web sites. Young people were represented by 10 names (e.g., Brittany, Kyle) obtained from Web sites listing contemporary baby names. These stimuli had been previously used by Mellott and Greenwald (1998) in their research on automatic ageism. From Bellezza et al. (1986), we selected 10 pleasant and 10 unpleasant words (e.g., paradise, poison) to represent the evaluative attribute. All names and evaluative words are listed in Appendix E.

Explicit measures. We used two feeling thermometers and five semantic differential scales to assess the favorability of participants' explicit attitudes toward old and young people. These were similar to the measures used in Experiment 1.

Design

The experimental design was a 2 (type of exemplars: pro-elderly vs. pro-young) \times 2 (IAT combinations: old + pleasant vs. young + pleasant) mixed factorial. The first factor was varied between subjects and the second within subjects. In addition, two other between-subjects counterbalancing factors were included: (a) order of IAT combinations (old + pleasant first vs. young + pleasant first) and (b) key assignment for exemplar classification (old with left key vs. young with left key).

Results and Discussion

IAT

As illustrated in Figure 2, exposure to pro-elderly exemplars yielded a substantially smaller automatic age bias effect (IAT effect = 182 ms, $d = 1.23$) than exposure to pro-young exemplars (IAT effect = 336 ms, $d = 1.75$), $F(1, 24) = 5.13$, $p = .03$. In addition, overall the IAT revealed significantly faster response latencies when pleasant stimuli were paired with young rather than old names (average IAT effect = 259 ms), $F(1, 24) = 101.69$, $p = .10^{-9}$.

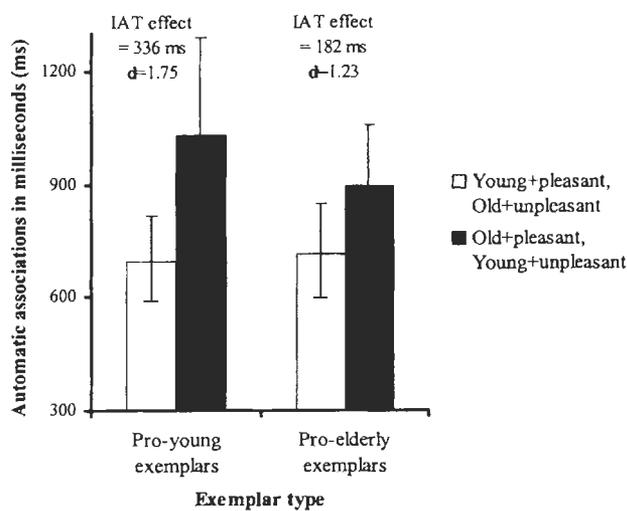


Figure 2. Effect of pro-elderly versus pro-young exemplars on implicit age associations. Error bars represent the standard deviation of responses in each condition. IAT = Implicit Association Test.

Simple effects showed that response latencies were faster for the old + pleasant IAT combination after exposure to pro-elderly than pro-young exemplars (895 vs. 1,028 ms, respectively). Latencies were also somewhat slower for the young + pleasant IAT combination after viewing the pro-elderly than pro-young exemplars (712 vs. 692 ms, respectively). However, neither of these effects reached statistical significance ($F_s < 1.96$, $p = ns$).

Feeling Thermometers

Feeling thermometers showed that participants' explicit attitudes toward young and old people were statistically similar ($M_s = 69^\circ$ and 68° , respectively; $F < 1$). Exposure to admired elderly exemplars did not produce greater liking for old people (difference between ratings of young vs. old people: $d = 0.40$ for the pro-young condition and $d = 0.34$ for the pro-elderly condition), Type of Exemplar \times Age Group: $F(1, 19) = 1.36$, $p = .26$.

Semantic Differential Scales

Unexpectedly, semantic differential scales revealed somewhat greater sensitivity to exemplar exposure. Participants preferred old over young people after exposure to admired elderly exemplars ($M_s = 1.40$ and 0.73 , respectively; $d = 0.83$), but expressed no preference for either group after exposure to admired young exemplars ($M_s = 0.74$ and 0.82 , respectively; $d = 0.09$), Type of Exemplar \times Age Group: $F(1, 19) = 3.43$, $p = .08$.

Relationship Between Implicit and Explicit Measures

For explicit measures, difference scores were computed by subtracting participants' ratings of old people from their ratings of young people such that higher scores indicated relative preference for young over old people. Correlations between explicit and implicit measures revealed a nonsignificant relationship between the IAT and feeling thermometers or semantic differential scales ($r = .32$, $p = .17$, and $r = .36$, $p = .11$, respectively). The two explicit measures were significantly correlated with one another, for both ratings of old ($r = .72$, $p < .0009$) and young people ($r = .70$, $p = .001$).

General Discussion

Although automatic attitudes have been previously conceptualized as relatively immutable (Bargh, 1999; Devine, 1989), the present research provides new evidence suggesting that automatic preference and prejudice may indeed be malleable. Two experiments demonstrated that implicit evaluations of historically stigmatized groups such as African Americans and older people may be modified, at least temporarily, by repeatedly reminding people of admired members of those groups and of disliked members of high-status reference groups.

The Malleability of Automatic Intergroup Attitudes

Experiment 1 showed that when famous African Americans and infamous European Americans were made salient, automatic race bias was reduced compared with the control and pro-White exemplar conditions. We conducted further analyses after combining data from Experiment 1 with an earlier pilot study that had used a

similar procedure.⁷ The combined analysis revealed that the IAT effect was reduced by more than half after exposure to pro-Black images (average IAT effect = 69 ms; $d = 0.53$) compared with pro-White images (average IAT effect = 191 ms; $d = 1.22$) and nonracial images⁸ (IAT effect = 174 ms; $d = 1.15$).⁹

The combined IAT data were analyzed further to identify the racial associations that were most influenced by counterattitudinal exemplars. These results indicated that the reduction of automatic White preference was primarily driven by faster latencies for Black + pleasant and White + unpleasant IAT combinations in the pro-Black condition than in the pro-White condition (mean response latencies = 739 ms vs. 868 ms, respectively), $F(1, 48) = 9.19, p = .004$. Latencies for White + pleasant and Black + unpleasant IAT combinations did not differ significantly across exemplar conditions (mean response latencies = 670 ms vs. 677 ms, respectively; $F < 1$). Thus the exemplar strategy appears to affect positive representations of African Americans or negative representations of European Americans or both, rendering them more accessible than before.

Our experiment relied on both admired Black and disliked White exemplars because we suspected that participants' implicit attitudes toward Blacks may be partially determined by their attitudes toward important reference groups, especially Whites. If so, exposure to exemplars from both groups may be important to elicit changes in chronic negative attitudes toward Blacks. However, for right now, this remains an open empirical question. Additional research is currently under way to determine if both admired Black and disliked White exemplars are necessary to moderate automatic race bias or if one type of exemplar is sufficient.

Although it is unclear whether implicit intergroup evaluations can be rendered malleable by admired versus disliked exemplars alone, it is clear that when presented together, they produce a substantial decrement in race bias that endures for 24 hr. We suspect that this effect is likely to fade after a longer hiatus; nevertheless, these findings are hopeful because they question the assumption that automatic racial attitudes are immutable because of their long socialization history. In addition, the strategy developed here may serve as a useful starting point in the attempt to produce more enduring changes in automatic prejudice and preference.

Experiment 2 showed that the influence of famous and infamous exemplars was not confined to the domain of racial attitudes but also readily applied to age-related attitudes. Automatic preference for younger over older people was also reduced after encountering images of admired older and disliked younger individuals. Together, these two experiments suggest that implicit intergroup biases may be modified at least temporarily with interventions that focus on changing the social context, that is, by creating environments that highlight admired and disliked members of various groups. Such experiences may, over time, render these exemplars chronically accessible so that they can consistently and automatically override preexisting biases.

Our data are consistent with a growing number of published and unpublished studies that have explored the malleability of automatic gender stereotyping and automatic racial stereotyping and prejudice using a variety of interventions (Blair, Ma, & Lenton, 2001; Haines, 1999; Kawakami et al., 2000; Lowery, Hardin, & Sinclair, 2001; Rudman, Ashmore, & Gary, 1999; Wittenbrink,

Judd, & Park, 2001). For instance, Blair, Ma, and Lenton (2001) found that a guided exercise imagining a strong woman decreased the ease with which perceivers automatically associated women with weakness and men with strength. In the domain of race prejudice, Rudman, Ashmore, and Gary (1999) found that students enrolled in a prejudice and conflict seminar expressed less automatic race prejudice than those enrolled in a race-unrelated class.

Potential Mechanisms Underlying Changes in Automatic Intergroup Evaluations

Two psychological processes may underlie the obtained shifts in implicit attitudes. If attitudes are conceptualized as context-dependent constructions (e.g., Smith, 1992; Smith & Zárate, 1992), the evaluation of an object (e.g., a social group) should depend on the subset of exemplars (individual members) retrieved from memory, which in turn should depend on exemplar accessibility. Any information stored with these exemplars, such as the perceiver's evaluation of them, is also likely to be applied to the judgment at hand. According to this model then, implicit evaluations of groups should be shaped by recently encountered and hence highly accessible stimuli (e.g., admired or disliked members) that are retrieved and applied to the judgment. If these exemplars become less accessible (e.g., in the absence of frequent exposure), their influence on group evaluations will recede and be replaced by other, perhaps default or chronically accessible, exemplars. In a nutshell, implicit evaluations of groups such as African Americans, European Americans, older people, or younger people depend on the context in which they are expressed and on the set of exemplars that are triggered in that context.

Alternatively, if attitudes are conceptualized as stable, stored evaluations (Jarvis, 1998; Petty & Jarvis, 1998; Wilson et al., 2000), exposure to liked members of a devalued group and disliked members of a valued group may create new abstract representations of target groups without erasing the old ones. These new knowledge structures may influence automatic attitudinal responses as long as they remain accessible. If the new representations become less accessible in memory, they should be less able to drive evaluative judgments.

The primary difference between the two processes described above is that the former suggests that attitude malleability depends on the subset of group-relevant exemplars temporarily activated in the social context and further suggests that no new abstract group representation need be created. The second mechanism argues that a new abstract representation of the target group is created that competes with the old representation. The first mechanism is more readily applicable to our research paradigm. Because the exemplars selected for our studies were famous and infamous individ-

⁷ The pilot study ($N = 17$) was our initial modest attempt to develop a racial exemplar strategy that could undercut the effect of automatic race preference. In this study, approximately half the participants viewed pro-Black images (admired Black and disliked White individuals), and remaining participants viewed pro-White images (admired White and disliked Black individuals); this study did not include a control condition. All participants then completed a race IAT and explicit attitude measures.

⁸ Data for this particular condition come from Experiment 1 only.

⁹ Participants' race and sex did not interact with any of the attitude measures in Experiments 1 and 2.

uals who were known to participants beforehand, it is unlikely that exposure to these exemplars over the course of the experiment elicited the creation of new mental representations. More likely, exposure simply increased the accessibility of known exemplars. Because our experiments did not directly pit these two mechanisms against one another, however, our choice of the underlying mechanism is clearly tentative.

Exemplar Exposure Appears Not to Change Explicit Intergroup Attitudes

We also found that the increased salience of pro-Black or pro-elderly images was insufficient to change self-reported evaluations of Blacks or older people as a group. The two distinct patterns of data captured by explicit versus implicit attitude measures are consistent with the small and nonsignificant correlations we found between explicit and implicit attitudes. We think that these data support the argument that implicit and explicit attitudes are at least partially dissociated and account for unique amounts of variance in people's responses (Dovidio et al., 1997; Fazio et al., 1995; Greenwald & Banaji, 1995; Wilson et al., 2000). Others who have reviewed the literature on the relationship between implicit and explicit attitudes (Blair, 2001; Devine & Monteith, 1999; Dovidio, Kawakami, & Beach, 2000) have offered both theoretical and methodological explanations for the weak association obtained on average, although all agree that a clear understanding of the implicit-explicit relationship continues to be elusive.

Our data also suggest that different processes may be responsible for changes in implicit versus explicit evaluations. In light of Martin's (1986) research on the correction of social judgments, one explanation for the explicit attitude finding is that having seen strongly positive and negative exemplars, participants concluded that these individuals were "exceptions to the rule" and corrected their attitudinal response before reporting it. This interpretation is also supported by a host of other findings demonstrating that explicitly held stereotypes can remain unchanged in the face of contradictory evidence (e.g., information about new members), provided that perceivers construe those new individuals as atypical (Bodenhausen, Schwarz, Bless, & Wanke, 1995, Experiment 3; Kunda & Oleson, 1995; Rothbart & John, 1985; Weber & Crocker, 1983). In such a situation, new cognitive categories (subtypes) are created to accommodate counterstereotypic individuals without changing the original stereotype. Indeed, subtyping and correction may be more evident when perceivers have the cognitive resources to reflect on and recategorize counterstereotypic exemplars but may be less likely to occur when attitude expressions are constrained by limited mental resources.

Implications for Changing Group Preference and Prejudice

Our data showing that exposure to admired and disliked group members produces substantial change in automatic intergroup evaluations but little change in explicit evaluations suggest that perhaps different types of strategies would be useful to combat automatic versus controlled expressions of prejudice and preference. This suggestion is reminiscent of another finding in the attitude literature demonstrating that attitudes formed by different processes (cognitive- vs. affect-based acquisition) require different

methods of change (Edwards, 1990; Edwards & von Hippel, 1995; Fabrigar & Petty, 1999). In that research, Edwards, Fabrigar, and their colleagues discovered that affect-based attitudes are changed more easily using persuasion tactics that rely on emotion, whereas cognition-based attitudes are changed more easily using persuasion tactics that rely on rational argument. If one applies similar logic to the domain of implicit and explicit attitudes, it is conceivable that whereas explicit attitudes may be best tackled with techniques that involve deep cognitive processing (e.g., increasing people's awareness of prejudice and persuading them to adopt egalitarian attitudes; see Monteith et al., 1993, 1994), implicit prejudice may benefit from the frequent use of techniques that involve shallower processing.

Implications for Media Influence on Automatic Attitudes

The mass media have been frequently criticized for disproportionately emphasizing stereotypic images of minorities and women (Greenberg & Brand, 1994; Harris, 1999). Interestingly, even when disliked members of dominant groups are portrayed in the popular media (e.g., news stories about White criminals like Jeffrey Dahmer), their race is typically not made salient. Rather, they are presented as deviant individuals, not members of a particular group. By contrast, news stories about Black criminals often highlight the individual's race. In our study, by forcing people to classify admired and disliked individuals by race, we emphasized the exemplars' group membership as well as their valence. These data imply that if media representations were to become more balanced, reminding people of both admired members of out-groups and less-than-stellar members of in-groups with emphasis on their group membership, the combined effect may be able to shift implicit prejudice and stereotypes.

In conclusion, we have provided new evidence for the malleability of automatic intergroup attitudes. Our goal is to offer a laboratory intervention with the hope that it will contribute to the development of future techniques capable of evoking more enduring changes in automatic intergroup attitudes.

References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin*, *71*, 319-342.
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2000). *Individual differences in physiological and self-reported measures of race-biased emotion*. Manuscript submitted for publication.
- Banaji, M. R., Hardin, C. D., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*, 272-281.
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY: Guilford.
- Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893-912.
- Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437-449.

- Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, and Computers*, *18*, 299–303.
- Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142–1163.
- Blair, I. V., Ma, J., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of automatic stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*, 828–841.
- Bodenhausen, G. V., Schwarz, N., Bless, H., & Wanke, M. (1995). Effects of atypical exemplars on racial beliefs: Enlightened racism or generalized appraisal? *Journal of Experimental Social Psychology*, *31*, 48–63.
- Brewer, M. B., & Miller, N. (1988). Contact and cooperation: When do they work? In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 315–326). New York: Plenum Press.
- Butler, R. N. (1980). Ageism: A foreword. *Journal of Social Issues*, *36*, 8–11.
- Chaiken, S., & Yates, S. (1985). Affective-cognitive consistency and thought-induced attitude polarization. *Journal of Personality and Social Psychology*, *49*, 1470–1481.
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.
- Cialdini, R. (1976). Elastic shifts of opinion: Determinants of direction and durability. *Journal of Personality and Social Psychology*, *34*, 663–672.
- Citrin, L. B., & Greenwald, A. G. (1998, April). *Measuring implicit cognition: Psychologists' and entomologists' attitudes toward insects*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.
- Cook, S. W. (1985). Experimenting on social issues: The case of school desegregation. *American Psychologist*, *40*, 452–460.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *121*, 163–170.
- Dasgupta, N. (2001). [Implicit attitudes toward lesbians and gay men and their relationship to perceivers' implicit sexual identification]. Unpublished raw data.
- Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: Ruling out the familiarity effect. *Journal of Experimental Social Psychology*, *36*, 316–328.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.
- Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 339–360). New York: Guilford.
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817–830.
- Dovidio, J. F., Kawakami, K., & Beach, K. R. (2000). *Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias*. Unpublished manuscript.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.
- Draine, S. C. (1998). *Inquisit* (Version 1.00) [Computer software]. Seattle, WA: Millisecond Software.
- Duckitt, J. (1992). Psychology and prejudice: A historical analysis and integrative framework. *American Psychologist*, *47*, 1182–1193.
- Edwards, K. (1990). The interplay of affect and cognition in attitude formation and change. *Journal of Personality and Social Psychology*, *59*, 202–216.
- Edwards, K., & von Hippel, W. (1995). Hearts and minds: The priority of affective versus cognitive factors in person perception. *Personality and Social Psychology Bulletin*, *21*, 996–1011.
- Fabrigar, L. R., & Petty, R. E. (1999). The role of the affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin*, *25*, 363–381.
- Farnham, S. D., & Greenwald, A. G. (1999, June). *Ingroup favoritism = implicit self-esteem × ingroup identification*. Paper presented at the annual meeting of the American Psychological Society, Denver, CO.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.
- Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*, 421–435.
- Fiske, S. T., & Neuberg, S. L. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*, 431–444.
- Forgas, J. P. (1992). On mood and peculiar people: Affect and person typicality in impression formation. *Journal of Personality and Social Psychology*, *62*, 863–875.
- Gaertner, S. L., & Dovidio, J. F. (1986). An aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). New York: Academic Press.
- Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology*, *77*, 669–687.
- Greenberg, B. S., & Brand, J. E. (1994). Minorities and the mass media: 1970s to 1990s. In J. Bryant & D. Zillman (Eds.), *Media effects: Advances in theory and research* (pp. 273–314). Hillsdale, NJ: Erlbaum.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Haines, E. L. (1999). *Elements of a social power schema: Gender standpoint, self-concept and experience*. Unpublished manuscript, City University of New York.
- Harding, J., Kutner, B., Proshansky, H., & Chein, I. (1954). Prejudice and ethnic relations. In G. Lindzey (Ed.), *The handbook of social psychology*. (Vol. 2, pp. 1021–1061). Cambridge, MA: Addison Wesley.
- Harris, R. J. (1999). *A cognitive psychology of mass communication* (3rd ed.). Mahwah, NJ: Erlbaum.
- Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.
- Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.
- Jarvis, W. B. G. (1998). *Do attitudes really change?* Unpublished manuscript, Ohio State University, Columbus.
- 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.
- Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality, and behavior. In J. F. Dovidio & S. L. Gaertner (Eds.), *Preju-*

- dice, discrimination, and racism* (pp. 35–60). New York: Academic Press.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effect of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888.
- Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology, 68*, 565–579.
- Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*, 275–287.
- Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology, 48*, 563–574.
- Lowery, B., Hardin, C. D., & Sinclair, S. (2001). Social tuning effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*, 842–855.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808–817.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33*, 471–489.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Wheeler, V. (1996). On resisting the temptation for simplification: Counterintentional effects of stereotype suppression on social memory. *Social Cognition, 14*, 1–20.
- Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology, 51*, 493–504.
- McConnell, A. R., & Leibold, J. M. (2001). Relations between the Implicit Association Test, explicit racial attitudes, and discriminatory behavior. *Journal of Experimental Social Psychology, 37*, 435–442.
- McFarland, C., Ross, M., & Conway, M. (1984). Self-persuasion and self-presentation as mediators of anticipatory attitude change. *Journal of Personality and Social Psychology, 46*, 529–540.
- McGuire, W. J., & Padawer-Singer, A. (1976). Trait salience in the spontaneous self-concept. *Journal of Personality and Social Psychology, 33*, 743–754.
- Mellott, D. S., & Greenwald, A. G. (1998, April). *Do older adults show automatic ageism?* Poster session presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.
- Mellott, D. S., & Greenwald, A. G. (1999, June). *Measuring implicit ageism: Comparing the Implicit Association Test and priming methods.* Poster session presented at the annual meeting of the American Psychological Society, Denver, CO.
- Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology, 65*, 469–485.
- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed vs. other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology, 64*, 198–210.
- Monteith, M. J., Zuwerink, J. R., & Devine, P. G. (1994). Prejudice and prejudice reduction: Classic challenges, contemporary approaches. In P. G. Devine, D. L. Hamilton, & T. M. Ostrom (Eds.), *Social cognition: Impact on social psychology*. San Diego, CA: Academic Press.
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167–184.
- Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition, 18*, 151–177.
- Myrdal, G. (1944). *An American dilemma: The Negro problem and modern democracy*. New York: Harper.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (1998, May). *Math = Bad + Male, Me = Good + Female, Therefore Math ? Me.* Paper presented at the annual meeting of the American Psychological Society, Washington, DC.
- Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effect of word familiarity and frequency on the Implicit Association Test. *Social Cognition, 19*, 97–144.
- Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology, 26*, 199–216.
- Petty, R. E., & Jarvis, W. B. G. (1998, October). *What happens to the "old" attitude when attitudes change?* Paper presented at the annual meeting of the Society of Experimental Social Psychology, Lexington, KY.
- Petty, R. E., Schumann, D. W., Richman, S. A., & Strathman, A. J. (1993). Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology, 64*, 5–20.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues, 41*, 81–104.
- Rudman, L. A., Ashmore, R. D., & Gary, M. (1999). *Implicit and explicit prejudice and stereotypes: A continuum model of intergroup orientation assessment.* Manuscript submitted for publication.
- Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437–465.
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*, 1315–1328.
- Sargent, M. J., & Theil, A. (2000). *When do implicit racial attitudes predict behavior? On the moderating role of attributional ambiguity.* Manuscript in preparation.
- Schwarz, N., & Bless, H. (1992). Scandals and the public's trust in politicians: Assimilation and contrast effects. *Personality and Social Psychology Bulletin, 18*, 574–579.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513–523.
- Sherman, S. J., Mackie, D. M., & Driscoll, D. M. (1990). Priming and the differential use of dimensions in evaluation. *Personality and Social Psychology Bulletin, 16*, 405–418.
- Smith, E. R. (1992). The role of exemplars in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment*. Hillsdale, NJ: Erlbaum.
- Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology, 71*, 888–898.
- Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3–21.
- Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660–1672.
- Stephan, W. G. (1985). Intergroup relations. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., Vol. 2, pp. 559–658). Hillsdale, NJ: Erlbaum.

- Strack, F., Schwarz, N., Bless, H., Kuebler, A. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology, 23*, 53–62.
- Swanson, J. E., Rudman, L. A., & Greenwald, A. G. (2001). Using the Implicit Association Test to investigate attitude-behavior consistency for stigmatized behavior. *Cognition & Emotion, 15*, 207–230.
- Tesser, A., & Danheiser, P. (1978). Anticipated relationship, salience of partner and attitude change. *Personality and Social Psychology Bulletin, 4*, 35–38.
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology, 45*, 961–977.
- Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 37–65). Hillsdale, NJ: Erlbaum.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*, 815–827.

Appendix A

Implicit Association Test Stimuli Used in Experiment 1

White names	Black names	Unpleasant words	Pleasant words
Josh	Lamar	Sickness	Rainbow
Andrew	Malik	Cancer	Gift
Brandon	Lionel	Vomit	Joy
Justin	Jamal	War	Paradise
Ian	Torrance	Poison	Laughter

Appendix B

Black and White Exemplars Used in Experiment 1

Name	True description	False description
Admired Black exemplars		
Martin Luther King	Leader of the Black Civil Rights movement in the 1960s.	Former Vice President of the United States.
Jesse Jackson	A longtime Civil Rights leader.	An actor famous for playing military characters.
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 <i>Fallen</i> .	Famous golf champion.
Eddie Murphy	Popular actor and comedian who was in the recent movie <i>The Nutty Professor</i> .	Famous American tennis player.
Michael Jordan	One of the world's best basketball players. Plays for the Chicago Bulls.	Lead singer in a popular rock band.
Tiger Woods	Professional golf champion.	Famous country music star.
Will Smith	Popular actor who was in <i>Men in Black</i> .	Famous TV talk show host.
Bill Cosby	Beloved actor and comedian who has his own TV show.	U.S. senator.
Gregory Hines	Famous actor and tap dancer who has his own TV show on CBS.	Newscaster for Dateline NBC.
Disliked White exemplars		
Ted Bundy	Serial rapist and killer from the Northwest.	Member of an international terrorist organization.
Jeffrey Dahmer	Serial killer who cannibalized his victims.	Bombed the World Trade Center in New York City.
Timothy McVeigh	Bombed the federal building in Oklahoma City killing hundreds.	Mafia member.
Charles Manson	Serial killer who operated in Los Angeles in the late 1960s.	Embezzled millions of taxpayers' money.
Al Capone	American gangster who terrorized Chicago in the 1920s.	Leader of an antigovernment militia.
Ted Kaczynski	The Unabomber who injured and killed using letter bombs.	Convicted pedophile.
Terry Nichols	Assisted in bombing the federal building in Oklahoma City killing hundreds.	Responsible for the crash of TWA Flight 800.
Howard Stern	Notoriously offensive radio talk show host.	Accused of embezzlement.
John Gotti	A high-ranking member of the Mafia.	Convicted for a series of bank robberies.
John Dillinger	Gangster in the Depression era who was a cold-blooded killer.	Responsible for bombing a Pan Am flight in the late 1980s.

(Appendixes continue)

Appendix B (continued)

Black and White Exemplars Used in Experiment 1

Name	True description	False description
Disliked Black exemplars		
O. J. Simpson	Charged with the brutal killing of his wife and her friend.	Embezzled millions of taxpayers' money.
Mike Tyson	Suspended from boxing for biting off his opponent's ear.	Convicted pedophile.
Louis Farrakhan	Leader of the Nation of Islam notorious for his anti-Semitic comments.	Bombed the World Trade Center in NYC.
Marion Barry	Former Mayor of Washington, DC, accused of buying and using cocaine.	Mafia member.
Arthur Washington	On the FBI's list of most wanted criminals. Member of militant Black prison groups.	Responsible for the crash of TWA Flight 800.
Lonny Gray	Member of the Denver Crips street gang.	Leader of an antigovernment militia.
Tyshawn Williams	Wanted by the FBI for murder.	Responsible for bombing a Pan Am flight in the 1980s.
Charles Brackett	Wanted by the FBI for murder and narcotics charges.	Member of an international terrorist organization.
Michael McClinton	Leader of a robbery ring.	Accused of embezzlement.
Stanley Obas	Accused of kidnapping and murdering a 13-year-old girl.	Convicted for a series of bank robberies.
Admired White exemplars		
Clint Eastwood	Actor famous for his roles in Westerns.	Famous country music singer.
Jim Carey	Actor and comedian who most recently acted in the <i>Truman Show</i> .	Famous TV talk show host.
Tom Cruise	Popular actor who starred in <i>Mission Impossible</i> .	Lead singer in a popular rock band.
David Duchovny	Acts as an FBI agent in the show <i>The X Files</i> .	Newscaster for Dateline NBC.
Tom Hanks	Received two Oscars for his acting roles in <i>Philadelphia</i> and <i>Forrest Gump</i> .	Famous American tennis player.
Jay Leno	Host of a popular late night TV show.	Famous golf champion.
John F. Kennedy	Former American President assassinated in Dallas.	Former Vice President of the United States.
Robert Redford	Famous actor and director.	U.S. Senator.
Norman Schwarzkopf	Retired General in the U.S. Army. Commander of Operation Desert Storm.	An actor famous for playing military characters.
Peter Jennings	News anchor for ABC.	U.S. Ambassador to the United Nations.

Appendix C

Flower and Insect Exemplars Used in Experiment 1

Flower stimuli	Insects stimuli
Carnation	Beetle
Daisy	Centipede
Hibiscus	Cockroach
Iris	Flea
Lily	Grasshopper
Orchid	Mosquito
Poppy	Scorpion
Rose	Spider
Sunflower	Tarantula
Tulip	Tick

Appendix D

Old and Young Exemplars Used in Experiment 2

Name	True description	False description
Admired elderly exemplars		
Barbara Walters	Co-anchor of a popular news show on ABC.	Award winning writer.
Mother Teresa	Was a true champion of the poor and destitute.	Famous novelist.
Jessica Tandy	Actress who won the Oscar for <i>Driving Miss Daisy</i> .	Famous singer.
Eleanor Roosevelt	Beloved former First Lady of the United States.	Award winning artist.
Frank Sinatra	Famous American singer.	Famous American painter.
Willie Nelson	Famous country music singer.	A popular D.J. on radio.
Albert Einstein	World famous physicist who won the Nobel Prize.	Popular jazz musician.
Phil Donahue	Popular talk show host.	Former tennis champion.
Dr. Benjamin Spock	Former pediatrician famous for his books on childrearing.	Famous golf champion.
Walter Cronkite	Former news anchor for CBS.	Famous actor.
Disliked young exemplars		
Diane Zamora	Army cadet who planned the brutal killing of her boyfriend's other lover.	School teacher who stole thousands from the teachers' retirement fund.
Tonya Harding	Olympic skater accused of plotting to injure her competitor in order to secure the Olympic gold for herself.	Daycare worker accused of neglecting children at her daycare center.
Susan Smith	Young mother in North Carolina who drowned both her children.	Convicted pedophile.
Louise Woodward	Nanny who was convicted of murder for killing the child in her care.	Convicted for dealing drugs.
Erik Menendez	Serving life in prison for killing his own parents with a shotgun.	Leader of an antigovernment militia.
Lyle Menendez	Brutally killed his own parents with a shotgun.	Convicted for a series of bank robberies.
Andrew Cunanan	Murderer who killed fashion designer Gianni Versace.	Embezzled millions of taxpayers' money.
Timothy McVeigh	Bombed the federal building in Oklahoma City killing hundreds.	Accused of insurance fraud.
Jeffrey Dahmer	Serial killer who cannibalized his victims.	Bombed the World Trade Center in New York City.
Ted Bundy	Serial rapist and killer from the Northwest.	Member of an international terrorist organization.
Disliked elderly exemplars		
Leona Helmsley	Infamous real estate tycoon convicted of felony and mail fraud.	School teacher who stole thousands from the teachers' retirement fund.
Marge Schott	Owner of the Cincinnati Reds team infamous for her racist and anti-Semitic comments.	Daycare worker accused of neglecting children at her daycare center.
Tammy Faye Bakker	Ex-wife of T.V. evangelist who helped him embezzle money from his church.	Convicted pedophile.
Marie Noe	Housewife accused of suffocating and killing 8 of her 10 children.	Accused of insurance fraud.
Marshall Applewhite	Leader of the Heaven's Gate cult who convinced his group to commit mass suicide.	Convicted for dealing drugs.
Ted Kaczynski	The Unabomber who injured and killed using letter bombs.	Leader of a large antigovernment militia.
James McDougal	Convicted of fraud and conspiracy in the recent Whitewater development case.	Recently convicted for a series of bank robberies.
Bob Packwood	Former senator who resigned after being accused of sexual assault and harassment.	Embezzled millions of taxpayers' money.
Sam Bowers	Former Ku Klux Klan head convicted for the brutal murder of a civil rights worker in Mississippi.	Bombed the World Trade Center in New York City.
Jim Bakker	Television evangelist convicted of embezzling millions of dollars from his church.	Member of an international terrorist organization.
Admired young exemplars		
Sarah McLachlan	Famous singer and organizer of the Lilith Fair.	Award-winning writer.
Jodie Foster	Oscar-winning actress.	Famous novelist.
Steffi Graf	One of the best female tennis players today.	Famous singer.
Princess Diana	Beloved princess who recently died in a car accident.	Award-winning artist.
Todd Eldridge	One of the finest American ice-skaters.	Famous American painter.
Leonardo DiCaprio	Actor who most recently played a leading role in <i>Titanic</i> .	A popular D.J. on radio.
Matt Damon	Popular actor who starred in <i>Good Will Hunting</i> .	Famous golf champion.
Ben Affleck	Popular actor and director.	Former tennis champion.
Casey Martin	Up-and-coming professional golfer who has succeeded despite his disability.	Famous actor.
Prince William	Heir to the British throne.	Up-and-coming golf champion.

(Appendixes continue)

Appendix E

Implicit Association Test Stimuli Used in Experiment 2

Young names	Old names	Unpleasant words	Pleasant words
Tiffany	Ethel	Sadness	Rainbow
Brittany	Agatha	Anger	Gift
Danielle	Bernice	Vomit	Joy
Christine	Lucille	War	Paradise
Julie	Agnes	Hell	Laughter
Justin	Oscar	Slum	Cuddle
Brandon	Alfred	Slime	Glory
Ryan	Clarence	Filth	Gold
Kyle	Cecil	Stink	Kindness
Corey	Irwin	Cockroach	Peace

Received February 14, 2000
Revision received December 26, 2000
Accepted January 2, 2001 ■



AMERICAN PSYCHOLOGICAL ASSOCIATION SUBSCRIPTION CLAIMS INFORMATION

Today's Date: _____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do **NOT** duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION _____ MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL) _____

ADDRESS _____ DATE YOUR ORDER WAS MAILED (OR PHONED) _____

CITY _____ STATE/COUNTRY _____ ZIP _____

YOUR NAME AND PHONE NUMBER _____ ISSUES: ___ MISSING ___ DAMAGED

TITLE _____ VOLUME OR YEAR _____ NUMBER OR MONTH _____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.

(TO BE FILLED OUT BY APA STAFF)

DATE RECEIVED: _____ DATE OF ACTION: _____
ACTION TAKEN: _____ INV. NO. & DATE: _____
STAFF NAME: _____ LABEL NO. & DATE: _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.