Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Experimental Social Psychology 41 (2005) 141–156
Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects☆

E. Ashby Plant*, B. Michelle Peruche, David A. Butz

Department of Psychology, Florida State University, Tallahassee, FL 32306-1270, USA

Received 17 September 2003; revised 26 May 2004
Available online 23 August 2004

Abstract

The current work proposes an approach for eliminating automatic bias by repeatedly exposing people to social stimuli where group membership (e.g., race) is unrelated to stereotypicality (e.g., being a violent criminal). Participants completed a computer program where they pretended they were police officers and decided as quickly as possible whether to shoot at Black and White suspects. Although initial responses to the program were biased by the race of the suspect, extensive practice with the program where race was unrelated to the presence or absence of a gun eliminated race biases immediately after practice (Study 1) and 24 h later (Study 2). However, this elimination of bias did not occur when race was related to the presence of a gun (Study 3). The final study (Study 4) revealed that extensive practice on the program led to the inhibition of racial concepts. The findings are discussed in terms of their implications for the elimination of automatic forms of bias.

Keywords: Implicit; Stereotyping; Person perception; Prejudice reduction

Introduction

The development of implicit measures of prejudice has allowed researchers to demonstrate the influence of racial bias on information processing by examining split second responses to stimuli (e.g., Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Wittenbrink, Judd, & Park, 1997). For example, many White people respond more quickly to negatively valenced words that are stereotypic of Black people when these words follow the presentation of Black faces as compared to White faces (Wittenbrink et al., 1997). Recently, much interest has focused on how these automatic biases can be altered with the hope of learning how to decrease these kinds of biases (see Blair, 2002 for a recent review; Devine, 2001). Several researchers have demonstrated that exposing people to counter-stereotypic stimuli can reduce bias (e.g., Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001). These findings are promising and indicate that influencing people's associations with social groups can reduce automatic forms of prejudice and stereotyping. The current work proposes an alternative, but related, approach to eliminating automatic forms of bias that repeatedly exposes people to social stimuli where the critical characteristic (i.e., being a violent criminal) is unrelated to group membership (i.e., race). The goal of this approach is to make race non-diagnostic and unhelpful for the decision making process and, thereby, encourage an unbiased response (e.g., Bodenhausen & Macrae, 1998).

Several recent studies examining changes in automatic bias have shown that altering the associations

☆ The authors gratefully acknowledge the assistance of Tyon Hall, Jeremy Ward, Dacia Stone, Jenn Dixon, Kathy Pennington, Sheba Paivandy, Michelle Lee, Faun Jimenez, Letrecia Spencer, Allison Clay, Jackie Steward, Kenisha Wilson, Patti Wallace, Nehemie Georges, and Karlista Peralta in collecting and entering the data for the samples reported in this manuscript.
* Corresponding author. Fax: +1 850 644 7739.
E-mail address: plant@psy.fsu.edu (E.A. Plant).

0022-1031/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2004.07.004

people have with social groups can reduce implicit stereotyping and prejudice. For example, Karpinski and Hilton (2001) demonstrated that when participants were exposed to word pairings that reversed young-good and old-bad associations (i.e., pairings of old-good and young-bad), participants' implicit biases about these groups decreased from previous levels. In addition, Dasgupta and Greenwald (2001, Study 1) found that exposure to admired Black and disliked White exemplars resulted in lower levels of implicit race bias than exposure to exemplars irrelevant to race. Blair et al. (2001) showed that counter-stereotypic mental imagery resulted in weaker implicit stereotype activation compared to neutral mental imagery or no mental imagery. In each of these studies, participants were exposed to counter-stereotypic information and this exposure altered the nature of their implicit responses.

The basic premise underlying these approaches is that by exposing people to counter-stereotypic information, the associations they have with social groups are altered, which in turn influences the characteristics or evaluations automatically activated upon exposure to group members (e.g., Blair et al., 2001; Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001). Specifically, by making associations with derogated groups more positive (or less stereotypic), exposure to these groups activates more positive (or less stereotypic) implicit responses. However, ideally, group membership (e.g., ethnicity, gender) would not influence responses at all. Such unbiased responses would seem possible if group membership were not helpful to people's decision-making process. For example, if people expect that Whites and Blacks are equally likely to be violent criminals, then the racial category (i.e., Black vs. White) of an individual is not diagnostic of criminality and, therefore, should not influence the decision about whether the individual is a violent criminal.

We were interested in whether it would be possible to eliminate an implicit racial bias by making group membership a non-diagnostic factor in responses to social stimuli. As opposed to only exposing people to information that runs counter to the stereotype, we explored whether multiple exposures to social stimuli where group membership (i.e., race) is statistically unrelated to the evaluated characteristic (i.e., being a violent criminal) would eliminate biased responses regarding this characteristic. In the absence of stereotypes and racial prejudice, people would ideally respond to individuals and evaluate them based on their personal characteristics without being biased by expectations (positive or negative) regarding the individuals' group membership. That is, repeatedly exposing people to stimuli where the critical characteristic is statistically unrelated to group membership may lead to unbiased responses.

In considering theoretical approaches to stereotyping, the proposed technique has the potential to decrease bias through two routes. First, because the proposed technique is designed in such a way that the critical characteristic (i.e., weapon possession) is statistically unrelated to group membership, group membership is not diagnostic. Therefore, attending to race will not aid and may instead impair performance on the task. Bodenhausen and Macrae (1998) proposed the existence of inhibitory mechanisms that "aid the perceiver in disregarding inappropriate, distracting, or interfering stimulus attributes in the process of person perception." (p. 9). As a result, because attending to race interferes with performance on the proposed task, the perceiver may come to disregard, and even inhibit, the racial category in the decision-making process (Bodenhausen & Macrae, 1998).

Second, similar to the processes believed to occur with exposure to only counter-stereotypic information, repeated and recent exposure to both counter-stereotypic and stereotypic stimuli could change the nature of category-based responses (Kashima, Woolcock, & Kashima, 2000; Kunda & Thagard, 1996; Smith & Zarate, 1992; Weber & Crocker, 1983). For example, if people begin with stereotype-consistent exemplars being more accessible than stereotype-inconsistent exemplars, viewing stimuli of both types in equal proportions should even out the relative accessibility of the two types of exemplars making them similarly accessible when making decisions about new social stimuli (e.g., Smith & DeCoster, 1998; Smith & Zarate, 1992). Alternatively, the inclusion of multiple counter-stereotypic stimuli could weaken the associative link between the group membership and the relevant characteristic (e.g., Kunda & Thagard, 1996).

Responses to criminal suspects

An unbiased, accurate response to individuals is not only vital to ensure fairness but in some cases may be a life or death matter. Consider the responses of police officers pursuing a criminal suspect. If a police officer possesses an expectation (i.e., stereotype) that Black people are more likely to be violent criminals than are White people, then split second decisions about whether or not a suspect is an imminent threat could be biased and lead to more aggressive responses to Black compared to White suspects. For example, when deciding whether or not to fire on a suspect, if police officers expect that Black people are more likely to be an imminent threat than White people, then this may influence how they interpret and respond to situations involving Black suspects with tragic consequences.

Consistent with this possibility, recent work indicates that the stereotype that Black people are more likely to be violent and criminal than White people (Devine & Elliot, 1995; Duncan, 1976) may influence the identification of weapons (Correll, Park, Judd, & Wittenbrink, 2002; Greenwald, Oakes, & Hoffman, 2003; Payne, 2001; Payne, Lambert, & Jacoby, 2002). For example, Correll et al. (2002) had participants decide quickly whether a male suspect who appeared on a computer screen possessed a gun or a neutral object (e.g., a cell phone). If the male suspect had a gun, they were instructed to shoot at the person by hitting a specified button. They found that when participants were forced to make decisions quickly (Study 2), they were more likely to mistakenly shoot at a Black person with a neutral object than fail to shoot at a Black person with a gun (also see Greenwald et al., 2003).

Given the types of biases found in participants' responses in these studies, it is critical to determine ways to eliminate race as a factor in determining responses. Exposure only to counter-stereotypic information (e.g., only unarmed Black suspects and armed White suspects) could create alternative biases (e.g., faster shooting of Whites and more mistaken shootings of Whites), which would be problematic in its own right. However, exposure to multiple decision trials where the race of the suspect is unrelated to the presence or absence of a gun could potentially eliminate biased responses for subsequent decisions without risking the creation of alternative biases. In order to explore this possibility, the current work examined whether repeated exposure to Black and White suspects who were equally likely to have a gun would eliminate the influence of race on future similar decisions.

In the current work, we examined whether non-Black participants' responses to criminal suspects were biased by the race of the suspect and whether extensive practice with the program where race was unrelated to the presence or absence of a gun (i.e., where race was non-diagnostic) eliminated race biases immediately after practice (Study 1) and 24 h later (Study 2). Next, we explored whether exposure to a program where race was related to the presence of a gun (i.e., diagnostic) would eradicate race bias (Study 3). Finally, Study 4 examined whether exposure to the program led participants to inhibit the activation of racial concepts and, thereby, eliminate the influence of race on responses.

Study 1

As a first step in examining these issues, Study 1 explored whether there were racial biases in responses to Black and White criminal suspects consistent with previous findings (e.g., Correll et al., 2002; Greenwald et al., 2003; Payne, 2001). In addition, Study 1 investigated whether training with a shooting task where race was unrelated to the presence of a gun would eliminate these biases. A computer program was developed where non-Black respondents were asked to pretend they were police officers and decide whether or not to shoot at suspects whose pictures appeared on the screen. The respondents were instructed to base their decision to shoot on whether a gun or some other object appeared in the picture. They were instructed to shoot at pictures where a gun was present by hitting a designated "shoot" key on the keyboard and not to shoot at pictures where some other object (e.g., cell phone, wallet) was present by hitting a designated "don't shoot" key. Half the pictures contained a gun and half contained some other object. In addition, half the pictures were of Black males and half were of White males. The program measured whether or not participants made the correct decision (i.e., shoot or don't shoot) based on the object present in the picture and the speed with which they made their decisions.

Because the program was specifically designed such that the faces were equally likely to be White or Black and each face was equally likely to be paired with a gun or neutral object, we anticipated that repeated exposure to the program would make race non-diagnostic and, therefore, unhelpful for participants' responses. This could occur either by leading participants to inhibit racial information when making their decisions (Bodenhausen & Macrae, 1998) or by changing the nature of the exemplars or associations activated when exposed to the Black and White faces (e.g., Kunda & Thagard, 1996; Smith & DeCoster, 1998; Smith & Zarate, 1992). In order to explore whether practice with the shooting task would eliminate participants' racial bias on the task, participants completed a large number of trials on the program, and we examined whether participants responded with racial biases on the early trials of the program and whether these biases were eliminated on the later trials. Because we did not want participants to artificially alter their responses to the Black and White faces in a manner that would decrease the bias (e.g., purposely bias their responses to hit the "don't shoot" key for the Black faces), the potential influence of race was not mentioned. Instead, participants were simply instructed to try to improve their performance across the trials.

Upon their initial exposure to the program, we anticipated that participants' responses would be biased by the race of the target picture. Specifically, because of the prevalent stereotype that Black people are violent criminals (e.g., Devine & Elliot, 1995) and based on recent findings on the misidentification of weapons (e.g., Correll et al., 2002; Greenwald et al., 2003; Payne, 2001) we expected that participants would be biased toward shooting when the face was Black compared to White. As a result, on the early trials of the program we anticipated that participants would be more likely to make errors when a Black face was paired with a neutral object (i.e., mistakenly shoot at a Black face with a neutral object) than when a Black face was paired with a gun. In contrast, when the face was White, participants were either expected to make the same number of errors regardless of whether a gun was present or to be biased away from shooting at White faces (Greenwald et al., 2003).

Further, the current study examined participants' responses to the simulation using the process dissociation procedure (e.g., Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993; Payne, 2001), which allows one to identify the distinct contributions of controlled and automatic processing to people's responses. When responding to the computer simulation, participants should, ideally, engage in controlled processing whereby they correctly distinguish guns from neutral objects. However, it is also possible that they will engage in more automatic processing where they are influenced by their stereotypic association between Black people and criminality, which may lead participants to shoot when the suspect is Black regardless of the object. Whereas controlled processes are consistent with participants' intentions and accurate performance on the task (i.e., correct responses to the simulation), automatic processes influence responses regardless of whether they aid performance on the task. In his work examining the influence of racial primes on the identification of weapons and tools, Payne found that the racial primes influenced the automatic processing of the information but not the controlled processing, such that participants had a larger automatic processing estimate for trials with Black compared to White primes.

As a result, in the current study, to the extent that participants' responses on the early trials of the program are influenced by an automatic racial bias, then the responses to the Black faces should reveal a larger automatic processing component than responses to the White faces. However, if the training decreases the influence of the automatic biases, then these differential degrees of automatic processing should be eliminated on the later trials. In such a case, the larger automatic component for the Black compared to White faces on the early trials should be eliminated for the later trials. Alternatively, in overcoming this bias, it is possible that participants would actively and consciously combat the bias by increasing their degree of control for the trials with Black faces. In such a case, participants would have a larger automatic component for the Black compared to White faces for both early and late trials (i.e., the automatic bias would not go away) but their controlled component would increase for the later trials with Black faces. That is, they would eliminate the bias by consciously working to improve the accuracy of their responses to Black suspects.

In examining responses to the computer simulation, we also investigated the speed with which the participants made their decisions (i.e., latency scores). It was possible that the latency scores would be influenced by the race of the face and the object paired with the face. For example, if participants expect that Black suspects are more likely to be armed, they might decide to shoot more quickly at Black faces paired with guns than White faces paired with guns. However, previous work has found that racial biases in latencies are more likely when participants have extensive time to make their decision. When participants are forced to make decisions quickly, racial biases tend to occur more in the number of errors than in the latency scores (Correll et al., 2002; Payne, 2001). Therefore, in the current study, we did not predict that race would influence the latency scores.

Methods

Participants and design

Participants were 125 non-Black introductory psychology students (70% female) who participated in exchange for course credit. Due to a problem with the computer program, age was not collected for 48 of the participants, but for the remaining participants the mean age was 19 years ($SD = .96$). Because all participants were drawn from the same participant pool, the age of the remaining 48 participants was likely similar to the other participants. Eighty-seven percent of the participants were White, 10% were Hispanic, 1% was Native American, and 2% were multi-racial. The experimental design was a 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) within-subjects factorial.

Materials

In order to test the current hypotheses, a computer simulation program was designed using Inquisit software. The computer program instructed participants: "Today you will be pretending that you are a police officer. Your task is to determine whether or not to shoot your gun. Pictures of people with objects will appear at various positions on the screen. Some of these people are dangerous criminals who have their guns drawn... Your goal is to determine as quickly as possible whether or not to shoot at the person in the picture. Some of the pictures will have a face of a person and a gun. These people are the criminals and you are supposed to shoot at these people. Some of the pictures will have a face of a person and some other object (e.g., a camera). These people are not the criminals and you should not shoot at them. Press the 'A' key for "SHOOT" press the '5' key on the keypad for "DON'T SHOOT"."

The program utilized digital color photographs of 9 Black and 9 White males with neutral facial expressions, which were selected from a set of slides (Malpass, Lavigueur, & Weldon, 1974) that had been matched for age and attractiveness. Three pictures of guns (2 black revolvers and 1 silver pistol) and 3 pictures of neutral objects (a camera, wallet, and cell phone) were formatted to be equivalent in size and similar in background color and contrast. One of these pictures (i.e., a gun or neutral object) was superimposed upon each of the faces in such a way that the face was still clearly visible (see Appendix for examples). However, where the object appeared on the face changed, so that the participant had to scan the face to find the image making it likely that participants were aware of the race of the face when responding. For each of the 18 faces 2 versions were created, one with a gun superimposed and one with a neutral object superimposed. This resulted in 36 pictures of 4 types: Black face with a gun, White face with a gun, Black face with a neutral object, and White face with a neutral object. The final pictures were four inches wide and six inches long.

The computer program randomly selected one of the 36 pictures and displayed the picture on the computer screen. In order to make the program sufficiently challenging, the picture randomly appeared toward the top, middle, or bottom of the screen and toward the right, center, or left of the screen. This ensured that the participant had to scan the full screen in order to find the image. Each picture appeared on the screen until the participant responded or until the 630 ms time limit was reached. Labels were pasted at the top of the screen reminding the participant of which keys to hit for shoot and don't shoot. Each participant completed 20 practice trials and then 160 test trials (2 sets of 80) each.

Participants were told that they would be completing many trials and their goal was to improve their performance (i.e., respond more quickly with fewer errors). As stated above, the potential influence of race on their performance was not mentioned so that the participants would not consciously alter their responses to show less bias. In order to ensure that errors on the program were sufficiently salient, when participants hit the wrong key or the time limit was reached, an error message appeared on the screen in red for a full second.

Procedure

The participants were run individually. They met the experimenter in the lab and were seated at the computer. After signing the consent form, the experimenter provided oral instructions regarding the computer program and then told the participant to begin the program. The program also provided instructions to ensure that the participant understood the procedure. After completing the program, participants were thanked, debriefed, and given their credit.

Results

We were interested in whether participants' performance on the program revealed less bias in the later trials than on the earlier trials. The trials were split in half and the responses to the first half of the trials were compared to responses on the later half of the trials. Responses on the early trials should reveal race bias such that participants should be biased toward shooting Black suspects. However, we anticipated that responses on the later trials would not reveal such a race bias. After analyzing the mean error rates and latency scores, the results were explored using a process-dissociation approach. Across the analyses, effects that are not explicitly mentioned were not significant.

Table 1
Number of errors as a function of trial, race of face, and object in Study 1

M (*SD*)	Error rates	
	White face	Black face
Early trials		
Gun	4.10(2.04)	3.20(2.02)
Neutral object	3.38(1.94)	3.95(2.40)
Late trials		
Gun	3.52(2.04)	2.93(2.11)
Neutral object	2.90(1.80)	2.58(1.77)

The error scores and latency scores were submitted to 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) repeated measures ANOVAs.[1] For the analyses, the response latencies were log-transformed. The statistics from the analysis on the log-latency scores are presented but for ease of interpretation, the untransformed mean latency scores are presented.

The analysis of error rates revealed a main effect of Trial such that participants made more errors on the early trials than the later trials, $F(1,124) = 41.79$, $p < .001$. See Table 1 for all means and standard deviations. This analysis also revealed a main effect of Race of Face such that, overall, participants made more errors on trials with White faces than trials with Black faces, $F(1,124) = 11.18$, $p < .002$. These main effects were qualified by a Trial by Object interaction, $F(1,124) = 7.15$, $p < .01$, and a Race of Face by Object interaction, $F(1,124) = 15.96$, $p < .001$. However, these lower-order interactions were qualified by a Race of Face by Object by Trial interaction, $F(1,124) = 10.58$, $p < .002$. To examine the nature of this interaction, the influence of Race of Face and Object were examined separately for the early and late trials.

Analyses of the early trials revealed a Race of Face by Object interaction, $F(1,124) = 23.37$, $p < .001$. When the face was White, participants made more errors when the object was a gun than when it was a neutral object, $t(1,124) = -3.35$, $p < .002$. However, when the face was Black, participants made more errors if it was paired with a neutral object than if it was paired with a gun, $t(1,124) = 3.31$, $p < .002$.

[1] It is worth noting that gender did not influence the findings of interest in any of the studies. As a result, gender is not included in the reported analyses.

Analysis of the error rates in the later trials revealed a main effect of Race of Face such that, consistent with the early trials, participants made more errors on the later trials with White faces than trials with Black faces, $F(1,124) = 10.50$, $p < .003$. The analysis also revealed a main effect of Object such that, consistent with the early trials, participants made more errors on trials with guns than on trials with neutral objects, $F(1,124) = 7.77$, $p < .007$. However, there was no interaction between Race of Face and Object. Of particular importance, on the later trials, participants responded with similar numbers of errors when Black faces were paired with neutral objects and Black faces were paired with guns, $t(124) = -1.61$, $p = .11$.

It is also worth noting that a comparison of the changes in error rates from the early to the late trials revealed that participants demonstrated the most improvement (i.e., decrease in error rates) for trials with Black faces with neutral objects (M improvement = 1.38), compared to all other types of trials (M improvement = .44), $t(1,124) = 3.76$, $p < .001$.

The analyses of the log-transformed latency scores revealed a main effect of Race such that participants responded more quickly to trials with Black faces ($M = 515$, $SD = 36$) than to trials with White faces ($M = 518$, $SD = 35$), $F(1,124) = 6.11$, $p < .02$. In addition, there was a main effect of Object, $F(1,124) = 20.34$, $p < .001$. However, this main effect was qualified by a Trial by Object interaction, $F(1,124) = 6.26$, $p < .02$. On the early trials, participants responded more quickly to trials with guns ($M = 518$, $SD = 36$) than to trials with neutral objects ($M = 521$, $SD = 38$), $F(1,124) = 26.67$, $p < .001$. On the later trials, this effect was far weaker, but participants continued to respond significantly more quickly when a gun was present ($M = 515$, $SD = 35$) than when a neutral object was present ($M = 519$, $SD = 35$), $F(1,124) = 4.87$, $p < .03$. It is important to note, however, that Race of Face did not interact with Trial either alone or in combination with Object, F's < 1.

Analyses using process dissociation approach. In evaluating participants' responses to the program, it is also useful to consider the process dissociation approach (Jacoby, 1991, 1993), which distinguishes between the contributions of controlled and automatic processing for responses. By using a set of algebraic equations provided by Jacoby (1991; also see Payne, 2001), the automatic and controlled components of participants' responses on the simulation task were isolated. For the current purposes, the automatic component (A) reflects the degree to which participants' responses are biased by the race of the suspect toward shooting. The controlled component (C) reflects the degree to which the participants are accurately categorizing the objects and responding appropriately to the program (i.e., shooting armed suspects only).

Table 2
Automatic and controlled estimates as a function of race of face and trial for Study 1

M (*SD*)	Race of face	
	White face	Black face
Automatic estimate		
Early trials	$.44^a$(.18)	$.55^b$(.22)
Late trials	$.45^a$(.24)	$.45^a$(.21)
Controlled estimate		
Early trials	.61(.16)	.63(.18)
Late trials	.67(.14)	.71(.15)

Note. For each the automatic estimate, means with differing superscripts differ significantly.

Participants' automatic processing estimates (i.e., A's) were submitted to a 2 (Race of Face: Black vs. White) × 2 (Trial: Early vs. Late) repeated measures ANOVA. This analysis revealed a main effect of Trial, $F(1,123) = 7.41$, $p < .008$, and a main effect of Race of Face, $F(1,123) = 10.03$, $p < .003$. However, these main effects were qualified by a Race of Face by Trial interaction, $F(1,123) = 9.19$, $p < .004$ (see Table 2 for all means and standard deviations). Planned comparisons revealed that, as predicted, participants responded with a larger automatic processing estimate when responding to Black faces on early trials than to White faces on early trials, $t(123) = 4.12$, $p < .001$. However, this difference did not approach significance for the later trials, $t(1,124) = .10$, $p = .93$. In addition, participants' responses to the Black faces on early trials had a stronger automatic component than responses to the Black faces on the later trials, $t(123) = 3.69$, $p < .001$. These findings indicate that automatic biases were influencing responses to Black faces for the early trials but the automatic influence was reduced after extensive exposure to the program.

The controlled processing estimates (C) were also submitted to a 2 (Race of Face: Black vs. White) × 2 (Trial: Early vs. Late) repeated measures ANOVA. There was a main effect of Trial, such that participants exerted more control over their responses on the late compared to early trials, $F(1,123) = 41.99$, $p < .001$. The analysis also revealed a main effect of Race of Face with participants exerting more control for their responses to Black compared to White faces, $F(1,123) = 11.61$, $p < .002$. There was no significant interaction between Race of Face and Trial, $F(1,123) = 1.97$, $p = .16$, indicating that exposure to the program was not influencing the control estimate differentially as a function of race.

Discussion

Participants' responses on the early trials of the program were consistent with predictions and previous work examining weapon identification (e.g., Correll et al., 2002; Greenwald et al., 2003; Payne, 2001). On the early trials participants made more errors when the Black faces were paired with neutral objects than when they were paired with guns. However, for White faces they made more errors when the White faces were paired with guns than with neutral objects. These findings indicate that participants' early responses to the computer program were biased by the race of the face in a manner consistent with the stereotype that Black people are more likely than White people to be violent criminals.

However, multiple exposures to Black and White faces where the race of the face was unrelated to the presence of a gun eliminated the biases in the errors to the program. That is, although on the early trials of the program, participants were more likely to mistakenly shoot at a Black suspect with a neutral object than mistakenly not shoot at a Black suspect with a gun, this bias was eliminated on the later trials. Further, examination of responses using a process dissociation approach indicated that the decrease in bias was due to participants reducing their biased, automatic processing when responding to Black faces from the early to the late trials. In addition, analysis of participants' controlled processing estimates indicated that that practice on the program had the added benefit of increasing the intentional control and, hence, accuracy of responses. Recent work by Greenwald et al. (2003) also found that accuracy on a similar shooting task improved over trials such that perceptual sensitivity was greater in the second half of trials compared to the first half of trials. However, they did not find that practice on their program influenced race bias, which may have been due to the fact that race was a relevant factor in their program. That is, in their program race was used as a cue to distinguish between police officers and criminal suspects.

Study 2

The fact that exposure to the program appeared to eradicate the biases evident on the early trials in Study 1 suggests that exposure to the computer task may be eliminating racial biases. However, it was important to examine whether the influence of exposure to the program would persist and influence responses at a later point in time. Therefore, we examined whether training on the computer program would influence responses 24 h after the initial training. In addition, in the current study, we also wanted to ensure that any elimination of bias was not due to other factors, such as familiarity with categorizing pictures or comfort in the experimental setting. Therefore, we included a control group, where participants at time 1 completed a computer task highly similar to the shoot/don't shoot program that required the categorization of pictures based on the content of the picture. Specifically, participants were shown pictures of flowers that contained either insects or birds and they were instructed to swat at the insects. This program was designed to familiarize them with a categorization task using similar responses on the same computer and in the same room as they would use at time 2. At time 2, all participants completed the shoot/don't shoot program.

To the extent that the bias elimination in the previous study was maintained over time, participants who trained on the shoot/don't shoot program at time 1 should continue to respond with a lack of race bias at time 2. Further, to the extent that the elimination in bias was due to exposure to the program where participants were repeatedly exposed to White and Black faces paired with guns and neutral objects, the elimination of bias should only be apparent for the participants who completed the shoot/don't shoot program at time 1. Therefore, participants who completed the swatting program with flowers and insects should respond with bias when exposed to the shooting program at time 2.

Methods

Participants and design

One hundred and twenty-three introductory psychology students (59% female, age $M = 18.64$) participated in exchange for course credit. Six additional participants completed the first session but did not return for the second session and three participants did not follow instructions on the computer program. These participants were not included in any analyses. Ninety-two percent of the participants were White, 2% were Hispanic, 1% were Asian, and 5% reported they were biracial, multi-racial, or their race was other. The experimental design was a 2 (Race of Face: Black vs. White) $\times$ 2 (Object: Gun vs. Neutral) $\times$ 2 (Condition: training vs. control) within-subjects factorial with Condition as a between-subjects factor.

Materials

To test the hypotheses, two computer programs were used in the current study. At time 1, half of the participants completed the shoot/don't shoot computer simulation used in the previous study (i.e., training group) where they imagined that they were a police officer and decided whether or not to shoot at suspects who appeared on the computer screen. The remaining participants at time 1 completed a computer task that required similar responses but in a domain unrelated to race where they were asked to imagine that they were a gardener and had to swat at insects that landed on their flowers (i.e., control group). If the picture of the flower that appeared on screen had an insect superimposed on it, they were directed to "swat" at the insect. However, if the picture had a bird on it, they were instructed not to swat. Participants were told to hit the 'A' key on the keyboard for "swat" and the '5' key on the numerical keypad for "don't swat". The participants were instructed to respond as quickly as possible as each picture appeared on the screen. Although the swat/don't swat program was clearly somewhat different in topic and nature of decision made than the shoot/don't shoot program, the swat/don't swat program did require identification of an object presented in the picture and a decision based on this object. At time 2, all participants completed the early trials of the shoot/don't shoot program.

Procedure

At time 1, the participants met the experimenter in the lab and completed the computer program individually. After signing the consent form, the experimenter provided oral instructions regarding the computer program (either shoot/don't shoot or swat/don't swat) and advised the participant that they could begin the program. The program also provided written instructions to insure that the participant understood the procedure. After completing the designated computer program, participants were thanked, given partial credit, and reminded that they were required to return 24 h later to complete the experiment.

Participants returned 24 h later to the same lab room, and at time 2, all participants completed the early trials of the shoot/don't shoot program. As in the first session, participants signed a consent form and were given oral instructions by the experimenter. After completing the program, participants were thanked, debriefed, and given their complete credit.

Results

In order to examine the current data, we examined responses to the early trials on the shoot/don't shoot program at time 2 for both participants who completed the shoot/don't shoot program at time 1 and those who completed the swat/don't swat program. Of interest was whether participants' responses at time 2 revealed racial bias as a function of the program completed at time 1.[2] If the training from time 1 continued to influence responses at time 2 for participants who had completed the shoot/don't shoot program, then responses at time 2 should not reveal race bias. Participants' number of errors made on the early trials of the program at time 2 were analyzed with 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Condition: training vs. control) mixed model ANOVA's with Race of Face and Object as the repeated factors.[3] All means and standard deviations are presented in Table 3. Analyses of the error scores revealed a marginal main effect of Condition, $F(1,112) = 3.10$, $p = .08$. Participants in the control condition made more errors overall than participants in the training condition. This analysis also revealed a main effect of Race of Face such that, overall, participants made more errors on trials with White faces than trials with Black faces, $F(1,112) = 8.19$, $p < .006$. In addition, there was a main effect of Object with participants making more errors on trials with guns than trials with neutral objects, $F(1,112) = 6.76$, $p < .02$. These main effects were qualified by a Race of Face by Object interaction, $F(1,112) = 10.13$, $p < .003$, which was in turn, qualified by a Race of Face by Object by Condition interaction, $F(1,112) = 5.08$, $p < .03$.

Table 3
Number of errors for early trials at time 2 as a function of condition, race of face, and object in Study 2

M (*SD*)	Error rates	
	White face	Black face
Training group		
Gun	3.82(1.98)	3.22(2.27)
Neutral object	3.03(2.01)	2.73(1.86)
Control group		
Gun	4.51(2.59)	3.19(2.13)
Neutral object	3.34(1.97)	3.77(2.26)

In order to explore the nature of this interaction, we conducted separate 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) within-subjects ANOVAs for the control and training groups. Responses in the control condition revealed a Race of Face by Object interaction, $F(1,46) = 11.10$, $p < .003$. When the face was White, participants made more errors when it was paired with a gun than when it was paired with a neutral object, $t(46) = 2.60$, $p < .02$. However, when the face was Black, participants made marginally more errors if it was paired with a neutral object than if it was paired with a gun, $t(1,46) = -1.81$, $p = .08$.

The analysis of the training condition revealed a main effect of Race of Face, such that participants responded with more errors when the face was White than when it

[2] Based on the findings from Study 1, training on the program at time 1 should result in the elimination of racial biases from early to late trials. Planned comparisons for participants who completed the training at time 1 revealed that when the face was Black, participants made more errors if it was paired with a neutral object ($M = 3.40$, $SD = 2.07$) than if it was paired with a gun ($M = 2.84$, $SD = 1.77$), $t(1,69) = -1.99$, $p = .05$. However, on the later trials participants responded with similar numbers of errors when Black faces were paired with neutral objects ($M = 2.53$, $SD = 1.61$) and guns ($M = 2.77$, $SD = 1.89$), $t(69) = 1.01$, $p = .32$.

[3] As in Study 1, analyses of the latency scores in this study and all subsequent studies resulted in no significant interactions involving Race of Face, all F's < 1. Therefore, the latency analyses are not presented.

Table 4
Automatic and controlled estimates as a function of race of face and condition for Study 2

M (SD)	Race of face	
	White face	Black face
Automatic estimate		
Training group	.44[a](.16)	.47[a](.18)
Control group	.45[a](.18)	.54[b](.16)
Controlled estimate		
Training group	.60(.16)	.65(.17)
Control group	.53(.17)	.57(.22)

Note. For each the automatic estimate, means with differing superscripts differ significantly.

was Black, $F(1,67) = 4.97$, $p < .03$. In addition, there was a main effect of Object, such that participants made more errors on trials with guns than on trials with neutral objects, $F(1,67) = 8.12$, $p < .007$.

Process dissociation analyses. Participants' automatic and controlled components at time 2 were analyzed with 2 (Race of Face: Black vs. White) × 2 (Condition: training vs. control) mixed model ANOVA's with Race of Face as the repeated factor. The analysis of the automatic component scores revealed a main effect of Race of Face such that, overall, participants were more accurate for trials with Black faces than White faces, $F(1,112) = 10.67$, $p < .001$. The predicted Race of Face × Condition interaction did not reach significance, $F(1,112) = 1.96$, $p = .16$. However, the planned comparisons were wholly consistent with predictions. All means and standard deviations are presented in Table 4. Participants in the control condition responded with a larger automatic processing component for trials with Black faces than for trials with White faces, $t(46) = 3.17$, $p < .004$. However, participants in the training program condition responded with similar automatic estimates for trials with Black and White faces, $t(67) = 1.43$, $p = .16$.

The analysis of the controlled estimates revealed a main effect of Race of Face with participants providing a higher controlled estimate for the Black trials than for White trials, $F(1,112) = 6.92$, $p < .02$. In addition, there was a main effect of condition with participants in the training condition responding with larger controlled processing estimates than participants in the control condition, $F(1,112) = 5.90$, $p < .02$.

Discussion

The findings from the current study were consistent with predictions. Participants that completed the shooting program at time 1 did not respond with race bias 24 h later. In contrast, participants who completed the insect swatting program at time 1 responded with a bias toward shooting Black faces on the early trials at time 2. These findings indicate that training with the shoot/don't shoot program eradicated biases up to 24 h after the initial training. In addition, they demonstrated that exposure to a similar task did not eliminate bias but that bias removal only occurred when training with the shoot/don't shoot program. Having further demonstrated the efficacy of practice with the shoot/don't shoot program, we next explored what aspects of the program were necessary for eliminating the racial bias. Specifically, we were interested in whether exposure to a program where race of face was related to the presence of a gun eradicated race biases.

Study 3

The previous studies demonstrated that extensive training on the shoot/don't shoot computer program led to the elimination of racial biases apparent in initial exposure to the program up to 24 h after training. We believed that the program was successful in large part because the Black and White faces were equally likely to be paired with a gun (i.e., race was non-diagnostic), which means that attending to racial information provided participants with no strategic advantage. However, it was also possible that training on the program led to the elimination in bias because participants were acquiring practice with the program, which allowed them to overcome their biases. In order to explore this possibility, a program was created where race of face was related to the presence of a gun. Specifically, in the revised program, Black faces were more likely to be paired with guns and White faces were more likely to be paired with neutral objects. If training on the revised program led to the elimination of race bias, then the elimination of bias apparent in the previous studies was likely due to practice. However, if exposure to the revised program did not eliminate race bias, then the bias elimination in the previous studies was likely due to exposure to a program where race was unrelated to the presence of a gun. It was also possible that exposure to the program where Black faces were more likely to be paired with guns would actually increase the race bias, because training on this program might increase people's expectations that Black faces would be paired with guns. However, to the extent that participants already expected that Black people were more likely to be criminals, exposure to the biased program should not alter responses.

Participants and design

Participants were 60 introductory psychology students (62% female, age $M = 19.30$) who participated in exchange for course credit. Eighty-five percent of the participants were White, 5% were Asian, 7% were Hispanic, and 3% were multi-racial or indicated their race was "other". The experimental design was a 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) within-subjects factorial.

Materials and procedure

The computer program used in Study 1 was modified in the current study. In the modified program the proportion of trials where Black and White faces were paired with neutral objects and guns were shifted. Specifically, for the trials with Black faces, 56 (i.e., 70%) contained a gun and 24 (i.e., 30%) contained a neutral object. In contrast, for the trials with White faces, 24 (i.e., 30%) contained a gun and 56 (i.e., 70%) contained a neutral object. In all other ways, the materials and procedure for the program were identical to Study 1.

Results

We were interested in whether participants' performance on the program revealed less bias in the later trials than in the earlier trials. As in Study 1, the trials were split in half and the responses to the first half of the trials were compared to responses on the later half. Because there were an unequal number of trials across the conditions, percentages were created for the number of errors on each type of trial (e.g., White/gun) by dividing the number of errors made on each type of trial by the total number of that type of trial completed. After analyzing the latencies and error rates, the process dissociation analyses are presented. Across the analyses, effects not explicitly mentioned were not significant.

The error scores were submitted to a 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) repeated measures ANOVA. This analysis revealed a main effect of Trial such that, overall, participants made more errors on the early trials ($M = .20$, $SD = .12$) than the later trials ($M = .16$, $SD = .12$), $F(1,59) = 21.07$, $p < .001$. This analysis also revealed a main effect of Race of Face such that, overall, participants made more errors on trials with White faces ($M = .19$, $SD = .12$) than trials with Black faces ($M = .17$, $SD = .12$), $F(1,59) = 6.25$, $p < .02$. These main effects were qualified by a Trial by Object interaction, $F(1,59) = 7.99$, $p < .007$, a Trial by Race of Face interaction, $F(1,59) = 4.27$, $p < .05$, and the predicted Race of Face by Object interaction, $F(1,59) = 11.41$, $p < .001$. The nature of the Race of Face by Object interaction indicated that across the early and the late trials, when the face was White, participants made more errors when the object was a gun ($M = .21$, $SD = .11$) than when it was a neutral object ($M = .17$, $SD = .09$), $t(59) = -2.31$, $p < .03$. However, when the face was Black, participants made more errors if it was paired with a neutral object ($M = .18$, $SD = .11$) than if it was paired with a gun ($M = .15$, $SD = .09$), $t(59) = 2.22$, $p < .03$. Importantly, the Race of Face by Object by Trial interaction did not approach significance, $F(1,59) < 1$, $p = .64$.

Process dissociation analyses. Participants' automatic and controlled processing estimates were submitted to a 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) repeated measures ANOVA. The analysis of the automatic estimates revealed a main effect of Trial, such that participants responded with larger A estimates on the early trials ($M = .53$, $SD = .14$) than the later trials ($M = .49$, $SD = .14$), $F(1,59) = 7.33$, $p < .01$. There was also a main effect of Race of Face such that participants had a larger automatic estimate for Black faces ($M = .54$, $SD = .14$) compared to White faces ($M = .48$, $SD = .14$), $F(1,59) = 12.67$, $p < .002$. The Race of Face by Trial interaction did not approach significance, $F(1,59) < 1$. In addition, on the later trials, participants' responses revealed a larger automatic estimate for the Black faces ($M = .52$, $SD = .15$) than the White faces ($M = .45$, $SD = .14$), $t(59) = 3.34$, $p < .002$. These findings indicate that unlike Studies 1 and 2, training on the program did not eliminate the stronger contribution of automatic processing for the Black compared to White faces.

Analyses of participants' controlled estimates revealed a main effect of Trial, such that participants responded with smaller control estimates on the early trials ($M = .40$, $SD = .21$) than the later trials ($M = .47$, $SD = .20$), $F(1,59) = 21.00$, $p < .001$. There was also a main effect of Race of Face such that participants had larger control estimates for Black faces ($M = .46$, $SD = .21$) compared to White faces ($M = .40$, $SD = .20$), $F(1,59) = 9.52$, $p < .004$.

Discussion

The findings from the current study indicate that the elimination of automatic bias apparent in the earlier studies after training on the shoot/don't shoot program where race was unrelated to the presence or absence of a weapon did not occur in the present study where race was related to the possession of a gun. Although practice on the program in the current study improved participants' accuracy, the automatic racial bias was not eliminated (see also Greenwald et al., 2003). That is, people's automatic processing estimate remained larger for Black faces than White faces even after training on the program. These findings indicate that it was not merely practice with the shoot/don't shoot program that eliminated the automatic race bias in the previous studies, but that it was likely the fact that race was unrelated to the presence of a gun. However, because the findings from the current study rely in part on a null finding (i.e., the racial bias was not eliminated), some caution should be taken when considering the findings. The final study examined whether the training program used in the prior studies eliminated bias because it led to the inhibition of racial category information.

Study 4

The previous studies demonstrated that the elimination in race bias following training on the shoot/don't shoot program lasted for 24 h and occurred when completing the shoot/don't shoot program where race of face was unrelated to the presence or absence of a gun but not when race was related to the presence of a gun. We next examined one potential reason why the training on the program decreases bias. Specifically, we posit that because the presence of a gun is unrelated to group membership in the program, race is not diagnostic and will not aid decisions on the program. Therefore, training on the program may lead the perceiver to disregard race on the task and perhaps, actually inhibit race. Bodenhausen and Macrae (1998) argued for the existence of inhibitory mechanisms that help the perceiver to ignore interfering information when forming impressions. It seems possible that over the course of multiple trials on the shooting task, participants will come to inhibit the target's race because of it's lack of predictive value and because attending to race actually impairs performance. If so, we would expect that participants who completed extensive training on the program would show inhibited activation of racial concepts compared to those who completed only a small number of trials. Alternatively, if the program eliminates bias by influencing the strength of the stereotype and/or the content of the information activated upon exposure to a Black person (e.g., Kashima et al., 2000; Kunda & Thagard, 1996; Smith & Zarate, 1992; Weber & Crocker, 1983), then exposure to the program should not influence the activation of racial concepts. That is, the category should still be activated even if the stereotype that is activated is changed or weakened (e.g., Lepore & Brown, 1997).

In order to examine these possibilities, participants either completed the full 160 trials of training or completed a small number of trials (i.e., 40). All participants were then asked to fill out a word completion task that included items that could be filled in with race-relevant words. To the extent that participants who completed the training on the program had come to inhibit race, they were expected to fill in fewer items with race-relevant words than those who had only completed a brief portion of the training. To provide an additional comparison, another group of participants only filled out the word completion task. This allowed us to determine whether training actually led to the inhibition of race-relevant concepts as compared to those who had not been exposed to Black and White faces in the program.

Method

Participants

Across the three conditions, 93 introductory psychology students (58% female, age $M = 19.38$, $SD = 1.82$) participated in exchange for course credit. Seventy-seven percent of the participants were White, 9% were Black, 1% was Asian, and 13% were Hispanic.[4]

Procedure and design

Sixty-four of the participants were randomly assigned to either the full training (long version, $n = 30$) or the short version ($n = 34$) of the shoot–don't shoot program. These participants met the experimenter in the lab individually and were seated at the computer. They were told that they would be completing two unrelated tasks—a shooting task on the computer for the Social Psychology lab and a word fragment completion measure for a Cognitive Psychology lab. After signing the consent form, the experimenter provided oral instructions regarding the computer program and then told the participants to begin the program. Participants completed either the long or short version of the shoot–don't shoot program used in Studies 1 and 2 where race was unrelated to weapon possession. Next, participants were administered the word completion task. Participants were given a maximum of 15 min to complete the word completion task, and then they were thanked, debriefed and given their credit.

An additional 29 participants were part of an appended control group that came into the lab in small groups and were given 15 min to fill out the word completion task. With the inclusion of the control group, the design of the study was a 3 (Program: long vs. short vs. none) between-subjects factorial.

Materials

The computer program was identical to the one used in Study 1. Participants in the long program condition completed the 20 practice trials and the full 160 trials of the shoot–don't shoot program. Participants in the short program condition completed the 20 practice trials and then 40 trials of the shoot–don't shoot program.

To measure the activation of racial concepts, participants were given a word completion task similar to those used by other researchers (e.g., Gilbert & Hixon, 1991; Sinclair & Kunda, 1999; Steele & Aronson,

[4] Black participants were included in the current study because Correll et al. (2002) demonstrated that race biases were apparent in their program for both White and Black participants. Careful examination of the results indicated that ethnicity did not influence responses across the outcome measures.

1995). However, as opposed to assessing stereotypic concepts, the task assessed the activation of racial concepts. Ten word fragments related to racial categories were provided (e.g., R__E, WH___, __ACK representing RACE, WHITE, and BLACK, respectively). These word fragments could also be filled in with non-race-relevant words (e.g., RULE, WHOLE, SMACK). The full list of race-relevant words were Black, minority, White, African, race, Harlem, ethnic, dark, racial, colored. Race-relevant word fragments were separated by 18 filler items (e.g., _OO_, _ASTE), which could only be completed with words not relevant to race (e.g., BOOK, WASTE). According to previous work (e.g., Gilbert & Hixon, 1991), when concepts are activated, people are more likely to fill in word fragments with words relevant to those concepts. Consistent with the procedure employed by Sinclair and Kunda (1999), participants were instructed to complete each word fragment with the first word that came to mind. Therefore the number of race-relevant word fragment completions served as a measure of the degree to which racial concepts were currently activated.

Results

The number of race-relevant words that participants filled in on the word completion task was compared across the conditions using a one-way 3 (Program: long vs. short vs. none) between-subjects ANOVA. The analysis revealed a significant main effect of program, $F(2,90) = 3.60$, $p < .04$. Post hoc tests revealed that participants who completed the full training completed significantly fewer race-relevant words ($M = 1.30$, $SD = .95$) than either participants who completed the short version of the program ($M = 2.00$, $SD = 1.15$) or those in the control condition who did not complete a version in the program ($M = 1.97$, $SD = 1.32$), both t's > 2.02, p's < .05. However, the participants who completed the short version and those who did not complete a version of the program responded with a similar number of race-relevant words, $t < 1$.

In addition, for the participants who completed the full training on the program, the error scores were submitted to a 2 (Race of Face: Black vs. White) × 2 (Object: Gun vs. Neutral) × 2 (Trial: Early vs. Late) repeated measures ANOVA. Although the 3-way Race of Face × Object × Trial did not reach significance, $F(1,24) = 1.08$, $p = .30$, planned comparisons indicated that the pattern of findings was highly consistent with the previous findings. Specifically, on the early trials, participants made more errors when Black faces were paired with neutral objects ($M = 5.80$, $SD = 2.98$) than when they were paired with guns ($M = 4.32$, $SD = 1.82$), $t(24) = -2.27$, $p < .04$. However, on late trials, they made a similar number of errors for Black faces paired with neutral objects ($M = 4.08$, $SD = 2.04$) and guns ($M = 3.88$, $SD = 1.51$), $t(24) = -.46$. $p = .64$. For the White faces, there was not a significant effect of Object for either the early or late trials, t's < 1.22, p's > .23.

To the extent that the inhibition of racial concepts apparent in the responses of the participants in the full training condition resulted in the decrease in their racial bias on the simulation, the number of race-relevant words should be associated with the degree to which participants decreased their racial bias on the simulation. In order to examine this possibility, a decrease in bias score was created by computing the degree to which participants made more errors when Black faces were paired with neutral objects compared to guns for the early and late trials. The bias score from the late trials was then subtracted from the bias score from the early trials to get an estimate of the size of the decrease in bias. The correlation between this decrease in bias score and the number of race-relevant words from the word completion task was $r = -.30$, $p = .14$, indicating that responding with fewer race-relevant words (and, hence, greater inhibition of racial concepts) was related to a greater decrease in the degree of racial bias on the computer simulation. This relationship suggests that the inhibition of racial concepts may be contributing to the decrease in racial bias on the shoot/don't shoot task. However, caution should be taken in drawing conclusions based on this relationship because the correlation did not reach conventional levels of significance. Because the correlation only focused on one cell of our design, we may have lacked sufficient power to find the effect. Further, the correlation may have been stronger if we had directly assessed the shift in activation of race-relevant words from the early trials to the later trials as opposed to only assessing activation at the end of the program.

Discussion

Whereas participants who completed the short version of the program or no program filled in similar numbers of race-relevant words on the word completion task (our measure of racial category activation), participants who completed the full course of 160 trials responded with fewer race-relevant words. These findings indicate that extensive practice with the program where race of face was unrelated to possession of a weapon led to the inhibition of racial concepts. Because race was non-diagnostic and paying attention to race only impaired performance on the shoot/don't shoot task, extensive exposure to the program encouraged the inhibition of the participants' racial categories.

Importantly, the inhibition of racial concepts apparent in our participants who completed training on the shoot/don't shoot program was unlikely the result of changes in the content or strength of their racial stereotypes. That is, although decreasing the activation of racial categories is likely to decrease the activation of stereotypes (i.e., perceivers can't draw upon a racial stereotype if they are not aware or not paying attention to the race of a target), a change in the content or strength of the stereotype would be unlikely to influence the overall activation of the racial category, at least in the short-term.

However, just because the program led to the inhibition of racial categories, it does not mean that the stereotype was not also influenced by exposure to the program. Instead, it indicates that something in addition to the change of the stereotype occurred with exposure to the program. Whereas previous bias reduction techniques have focused on the effects of training using counter-stereotypic information to directly influence the activation of stereotypes (e.g., Dasgupta & Greenwald, 2001; Kawakami, Dovidio, Moll, Hersen, & Russin, 2000), the present approach appears to encourage the inhibition of the racial category, thereby, leading participants to disregard the race of the suspects in their responses. Such inhibition of racial categories should have the added benefit of diminishing the activation and application of any racial stereotypes.

General discussion

The current work examined whether racial biases in responses to criminal suspects could be eliminated by repeated exposure to suspects where the race of the suspect was unrelated to the presence of a gun. Consistent with previous work (e.g., Correll et al., 2002; Greenwald et al., 2003; Payne, 2001), participants' responses to the early trials of the shooting task were biased by the race of the suspect in a manner consistent with the stereotype that Black people are violent and criminal. That is, when the face was Black, participants were more likely to mistakenly shoot at a face paired with a neutral object than mistakenly not shoot at a face paired with a gun. In contrast, when the face was White, participants were more likely to make errors when the White faces were paired with guns compared to neutral objects.

These findings are highly troubling and suggest that responses to criminal suspects may be biased by the race of the suspect. Such biases, if present among police officers, could lead to tragic outcomes. For example, in situations where police officers must decide quickly if a person is an imminent threat, they may be more likely to mistakenly shoot at an unarmed Black person than at an unarmed White person. Of course, it is a far reach to go from the responses of undergraduate students on a computer program to decisions made in the field by police officers, but such possibilities exist and need to be explored.

On a more promising note, in the current work, training with the shooting program where the race of the face was unrelated to the presence or absence of a gun decreased the types of racial biases evident in early responses to the program immediately after training (Study 1) and 24 h after initial training (Study 2). That is, although participants responded with racial bias in their error rates on the early trials of the program, these biases were eliminated in the later trials of the program and 24 h later. Process dissociation analyses indicated the elimination of bias was not merely due to a shift in their overall control over responses but resulted due to a decrease in the automatic component (i.e., biased tendency) in participants' responses to Black faces. These findings are heartening and suggest that such bias is not inevitable. In addition, the efficacy of the current approach indicates that in order to eliminate implicit racial bias, it is not necessary to expose people only to stereotype-inconsistent information (e.g., Blair et al., 2001; Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001). Instead, presenting people with equal amounts of stereotype consistent and inconsistent information and, thereby, making race non-diagnostic, may eliminate the influence of race on subsequent responses.

In addition, examination of the findings from the current work using the process dissociation approach provides some insight into the aspects of participants' responses that were influenced by the training. Although training on the program increased the degree to which participants' exerted intentional control over their responses (i.e., C scores increased with training), this was true for both trials with Black and White faces. As a result, the increased control could not account for the decrease in bias. However, training eliminated the heightened automatic, unintentional processing component for the Black trials as compared to the White trails. These findings indicate that training directly influenced the degree of automatic racial bias as opposed to resulting in some degree of controlled, conscious compensation for the bias.

In considering why the training was effective, the findings from Study 3 indicate that practice using the program cannot account for the elimination of racial bias. When race was related to the presence of a gun, the bias was not eliminated in later trials. Importantly, although in Study 3 when the race of face was related to the presence of a gun participants' control over their responses improved between early and late trials of the program, there was no differential shift in their automatic processing component for the Black compared to White faces.

The findings from Study 4 revealed that repeated exposure to trials where race was non-diagnostic led to inhibition of racial concepts. Because the presence of a gun is unrelated to group membership and is not a helpful indicator for the decisions made in the program, repeated exposure to the program likely led to the inhibition of race. These responses are consistent with Bodenhausen and Macrae's (1998) proposition that people will keep distracting or extraneous information from influencing their ability to draw upon task-relevant information. On the current simulation task, this could be conceptualized as people keeping the race of face from influencing their ability to discern threatening from non-threatening objects. We believe that, over the course of multiple trials on the shooting task, because race was not an informative cue, participants came to inhibit the activation of the racial category and, thereby, eliminated the automatic influence of race on their responses.

At the outset of this paper we raised the alternative possibility that repeated and recent exposure to an equal number of Black people who are stereotype inconsistent (i.e., unarmed) and stereotype consistent (i.e., armed) could influence the strength or content of the stereotype of Black people (e.g., Kashima et al., 2000; Kunda & Thagard, 1996; Smith & Zarate, 1992; Weber & Crocker, 1983). For example, based on Smith and his colleagues' work on exemplars (e.g., Smith & DeCoster, 1998; Smith & Zarate, 1992), exemplars that have been frequently and recently encountered are likely to be highly accessible and influence later responses to similar targets. Although the fact that exposure to the program led to the inhibition of racial concepts indicates that, at least in part, participants were coming to ignore or disregard racial category information, it is also possible that exposure to the program influenced the nature of the stereotype.

Recently Payne et al. (2002) examined the potential bias-eliminating effects of explicitly manipulating attention to race by instructing some participants to avoid the use of race. Interestingly, compared to a no instruction control group, those told to avoid race actually committed more stereotype-consistent errors. Payne and colleagues concluded that explicit instructions to avoid race may result in more stereotypic responses, and, therefore, may be an ineffective bias-reduction approach. The current approach, however, achieves avoidance of race without explicit instructions. Rather than informing participants that race should not be attended to, race was a non-diagnostic cue and participants were given multiple instances with which to reinforce the non-diagnosticity of race. It is likely that allowing participants to "discover" that race was not a helpful cue actually aided the elimination of bias. If we had instead informed people that race was non-diagnostic, they may have shown increased bias compared to those not informed, findings similar to those of Payne and his colleagues.

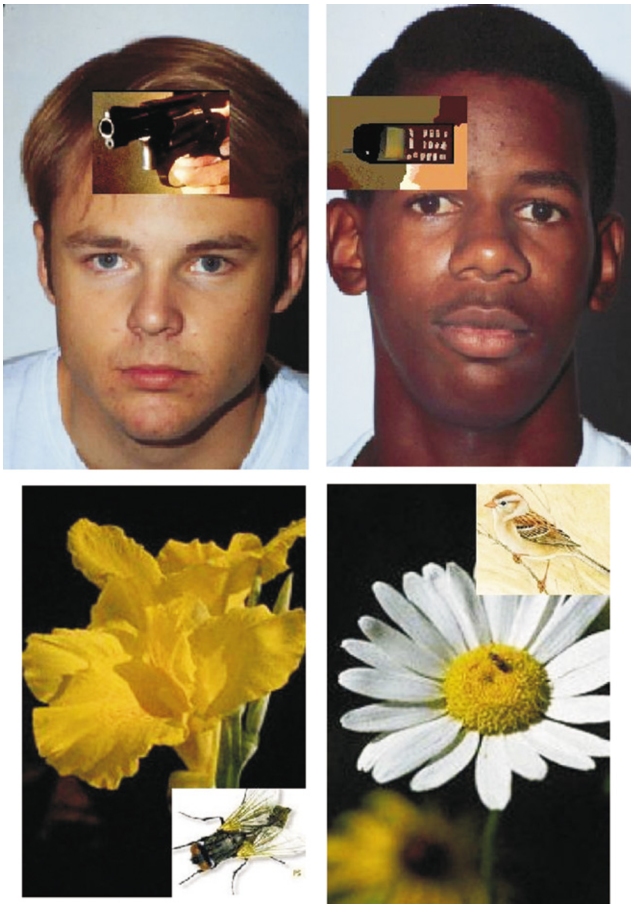
In the current work, the effectiveness of the program was only examined for responses up to 24 h following the initial training (also see Dasgupta & Greenwald, 2001; Kawakami et al., 2000). Practically, an effective bias elimination procedure should influence responses long after the initial procedure. In future work, it will be important to investigate whether the effects of the current approach to bias elimination last beyond 24 h after the initial training. It will also be informative to examine whether there are individual differences that influence the degree of racial bias revealed in the initial exposure to the shooting task as well as the effectiveness of the shooting task. For example, one could imagine that training on the shooting task may be particularly effective for individuals who strongly endorse the stereotype that Black people are violent criminals and are particularly likely to respond with bias on initial trials of the shooting task (see Correll et al., 2002).

It seems quite possible that the current approach would be effective for eliminating other types of automatic biases as well. Consider, for example, Karpinski and Hilton's (2001) approach where participants were exposed to word pairings that reversed young-good and old-bad associations (i.e., parings of old-good and young-bad). Repeated exposure to pairings where old and young were equally likely to be matched with good and bad may decrease the implicit biases about these groups. Such an approach would have the added benefit of avoiding the potential reversal of the prior evaluation bias (i.e., judging young as bad and old as good). It is also possible that the current approach could be used for training of police officers in order to eliminate racial biases in responses to criminal suspects. If effective, this relatively simple training technique could help to eliminate tragic mistakes.

Conclusions

The current work demonstrated the efficacy of a new approach to bias elimination that, as opposed to only exposing people to information that runs counter to the stereotype, exposed people to social stimuli where group membership (i.e., race) was statistically unrelated to the evaluated characteristic (i.e., being a violent criminal). Instead of only shifting the nature of people's stereotypic response, the current approach led to the inhibition of the racial category and eliminated their automatic racial biases. As we argued at the outset of this paper, ideally, people should respond to individuals based on their personal characteristics without being biased by their social group membership. We believe that the current approach holds promise for encouraging responses that are not influenced by automatic, stereotype-based expectations.

Appendix A. Example pictures from programs



References

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*(3), 242–261.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 5*(81), 828–841.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer (Ed.), *Stereotype activation and inhibition* (pp. 243–256). Mahwah, NJ: Erlbaum.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 86*, 1314–1329.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800–814.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5–18.

Devine, P. G. (Ed.), (2001). Implicit prejudice and stereotyping: How automatic are they? [Special section]. *Journal of Personality and Social Psychology, 81*(5).

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 21*(11), 1139–1150.

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*(5), 510–540.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping Blacks. *Journal of Personality and Social Psychology, 34*(4), 590–598.

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*(6), 1013–1027.

Gilbert, D. T., & Hixon, G. J. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509–517.

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*(6), 1464–1480.

Greenwald, A. G., Oakes, M. A., & Hoffman, H. G. (2003). Targets of discrimination: Effects of race on responses to weapons holders. *Journal of Experimental Social Psychology, 39*, 399–405.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139–154.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology, 81*(5), 774–788.

Kashima, Y., Woolcock, J., & Kashima, E. S. (2000). Group impressions as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review, 107*(4), 914–942.

Kawakami, K., Dovidio, J. F., Moll, J., Hersen, S., & Russin, A. (2000). Just say no to stereotyping: Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*(5), 871–888.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*(2), 284–308.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*(2), 275–287.

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1974). Verbal and visual training in facial recognition. *Perception and Psychophysics, 14*(2), 285–292.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*(2), 181–192.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38*(4), 384–396.

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology, 77*(5), 885–904.

Smith, E. R., & DeCoster, J. (1998). Knowledge acquisition, accessibility, and use with a recurrent connectionist network. *Journal of Personality and Social Psychology, 74*(1), 21–35.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgement. *Psychological Review, 99*, 3–21.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology, 45*(5), 961–977.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*(2), 262–274.