In recent years, there have been multiple occasions in the United States where police have shot and killed unarmed Black men after reportedly thinking the suspect was armed. Salient examples include the fatal shootings of Amadou Diallo, Anthony Dwain Lee, and Timothy Thomas. Although police occasionally shoot unarmed White suspects, too, observers have wondered whether police are more likely to use lethal force with Black suspects (Cardwell & Chan, 2006). Some have questioned the ability of officers to perform their duties impartially, raising doubts about police training (Fernandez, 2008).

Social psychologists have investigated racial bias in the decision to shoot as well as the effect of training. In much of this work, researchers attempt to recreate the predicament of an officer confronted with a potentially hostile suspect (e.g., Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007; Plant & Peruche, 2005). Typically, the race of the suspect and the object in possession are manipulated using a first-person-shooter task (FPST). In one such task, participants view male targets who are either Black or White, and holding either a gun or a nongun object (e.g., a wallet; Correll, Park, Judd, & Wittenbrink, 2002). When an armed target appears, participants must press a button to indicate shoot. When an unarmed target appears, participants must press a different button to indicate don’t shoot.

The race of the target in this paradigm typically has pronounced effects. Undergraduate participants are faster to shoot armed targets when they are Black (rather than White), and they are faster to choose don’t shoot for unarmed targets when they are White (rather than Black). A similar bias emerges with error rates. Participants are more likely to mistakenly shoot unarmed Black targets and to mistakenly choose don’t shoot for armed Whites. Signal detection theory (SDT) sheds light on these errors. SDT simultaneously estimates two parameters. One parameter, called $d’$ or sensitivity, represents participants’ ability to differentiate armed from unarmed targets. The other parameter, called $c$ or criterion,
represents participants’ willingness to open fire. A low or lenient criterion suggests that participants fire frequently; a high criterion suggests that participants fire rarely. Across multiple studies, we have found that sensitivity does not depend on race—Participants are capable of distinguishing between armed and unarmed targets, regardless of whether those targets are Black or White. However, participants show pronounced differences in the criteria they use: They use a much lower criterion when a target is Black rather than White and thus shoot more Black targets overall (Correll et al., 2002; Greenwald, Oakes, & Hoffman, 2003; Payne, 2001).

The Role of Stereotype Accessibility

This pattern of bias seems to reflect the accessibility of stereotypes that link Blacks (particularly young Black men) to the concept of danger (Correll et al., 2002; Correll, Park, Judd, & Wittenbrink, 2007; Devine & Elliot, 1995). Because past work bears directly on the current research, we review these studies in detail.

Manipulating Stereotypes via Salient Exemplars

In one study, participants read newspaper articles describing a series of violent crimes (Correll, Park, Judd, & Wittenbrink, 2007, Study 1). Some read articles that described the suspects as Black men; others read articles that described the suspects as White men. After reading the reports and viewing police sketches of the suspects, participants performed a FPST. Participants who read about Black criminals showed more extreme bias, adopting a more lenient or “trigger-happy” criterion for Black targets and a relatively conservative criterion for Whites. But when experimenters undermined the Black-danger association by exposing participants to articles about White criminals, participants showed no bias whatsoever.

Manipulating Stereotypes via Base Rates

In a conceptual replication, the authors manipulated stereotype accessibility within the FPST by changing the base rates of Black and White targets who were armed versus unarmed (Correll, Park, Judd, & Wittenbrink, 2007, Study 2). In a preliminary round, participants in one condition performed a task that reinforced the Black-danger association. This task presented 20 armed Black targets but only 12 armed White targets, and 20 unarmed White targets but only 12 unarmed Black targets. In essence, the task created an environment congruent with cultural stereotypes: Blacks were associated with danger. Participants in another condition performed a task that undermined the cultural stereotype by presenting mostly unarmed Black targets and armed White targets, creating a stereotype-incongruent (SI) association between Whites and guns. After exposure to these various environments, participants completed a FPST in which target race and object were unrelated. Participants who had been exposed to predominantly stereotype-congruent (SC) targets in the preliminary round showed greater bias in response times than participants exposed to predominantly SI targets.

Together, the research suggests that stereotype accessibility affects racial bias on the FPST. Two distinct experimental manipulations—one relying on salient exemplars and one relying on base rates—reinforced (or undermined) the association between Blacks and danger, and as a result, exacerbated (or attenuated) racial bias in the decision to shoot.

Training as a Buffer

When police officers complete the standard FPST, they—like civilians—respond more quickly to targets who conform to the Black-danger stereotype (i.e., armed Blacks, unarmed Whites) and more slowly to targets who defy the stereotype (i.e., unarmed Blacks, armed Whites). The observed bias in reaction time suggests that the targets are triggering stereotypes. But, strikingly, officers show no evidence of racial bias in their ultimate decisions to shoot, as measured by SDT criteria. In other words, although their response times show bias, suggesting that officers activate the Black-danger stereotype, their ultimate decisions reveal no bias (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). This pattern suggests that, in executing decisions to shoot, officers are able to either (a) effortfully override the stereotypes that influence their reaction times or (b) base their decisions on task-relevant diagnostic information (i.e., the object in the target’s hand) instead of race.

The present study is concerned with this capacity to respond without bias and how it develops through training. We suggest that training and expertise may help officers overcome or ignore the stereotypes that generally affect decisions made by people who lack the training. Indeed, an investigation of lab-based training found that undergraduate participants who were given extensive practice on the FPST showed patterns of performance that mirrored police: They showed reduced bias in their decisions to shoot even though they continued to display bias in their response times (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). Although lab-based training hardly simulates a veteran officer’s training and on-the-job experience, it provides some indication that training may attenuate the influence of stereotypes.

Overview of the Present Research

The present research attempted to shed light on the processes by which training and expertise reduce racial bias in the decision to shoot. We were guided by two hypotheses. First, if training and expertise truly enable experts to overcome
or ignore stereotypes when executing their responses, experts should be insensitive to manipulations of stereotype accessibility. When Black-danger stereotypes are highly accessible, novices should respond with greater bias (replicating Correll, Park, Judd, & Wittenbrink, 2007), but experts should be unaffected (Hypothesis 1). Such a finding would suggest that training enables participants to overcome or ignore the influence of race and attend instead to the presence or absence of a weapon.

This perspective also suggests possible boundary conditions, in that training and expertise should only reduce bias to the extent that participants are forced to attend to diagnostic object information (i.e., gun vs. no gun) instead of race. Our second hypothesis addressed this particular issue. When training and experience, themselves, reinforce the association between Blacks and danger, they should not reduce bias (Hypothesis 2).

Across three studies, we attempted to manipulate orthogonally (a) the accessibility of stereotypes linking Blacks with danger and (b) the expertise of participants (using both experimental, lab-based manipulations of training as well as known groups, comparing untrained participants to sworn law enforcement officers).

**Experiment 1**

We sought to test the hypothesis that police officers (who have law-enforcement training and experience) and “expert” undergraduates (who acquire training in the lab) would be less affected by manipulations of stereotype accessibility than “novice” undergraduates (who do not practice the FPST). Building on earlier research, we asked participants to read newspaper articles that described violent criminals as either Black or White. We then tested the impact of this accessibility manipulation on the performance of three groups: novice undergraduates, expert undergraduates, and sworn officers.

**Method**

**Participants and Design.** A total of 75 undergraduates (43 female, 32 male; 54 White, 11 Asian, 7 Black, 3 Latino/a; M age = 20.13 years) participated for either course credit or US$10. Police were also recruited for this study. Participation was completely voluntary, and officers were assured that there would be no way to identify individual performance on the task. Fifty-two officers completed the study during off-duty hours for US$45 (4 female, 48 male; 40 White, 7 Black, 5 Latino/a; M age = 39.75 years). The reported results are based on the performance of non-Black participants.1 Thus, the final sample of 113 participants included 45 police officers and 68 undergraduates (41 female, 72 male; 94 White, 11 Asian, 8 Latino/a).

Undergraduates were randomly assigned to one of two training conditions (novice or expert), which—with the addition of the officers—yielded three groups varying in training and expertise. All participants were then randomly assigned to read one of two fabricated newspaper articles (Black-criminal or White-criminal) before performing the FPST. Thus, the experiment involved a 3 (Training Condition: novice undergraduate vs. expert undergraduate vs. police officer) × 2 (Article Condition: Black-criminal vs. White-criminal) × 2 (Target Race: Black vs. White) × 2 (Object Type: gun vs. no gun) mixed-model design, with repeated measures on the last two factors (representing the FPST).

**Materials**

**FPST.** The FPST presented Black and White male targets embedded in unpopulated background scenes. Targets were holding either a gun or an innocuous object (see Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007, for details). Using either a button box or keyboard, participants were instructed to press a button labeled shoot if the target was armed and to press a button labeled don’t shoot if the target was unarmed. The task awarded points based on performance. Correctly indicating don’t shoot in response to an unarmed target earned 5 points (correct rejection), but indicating shoot earned a penalty of 20 points (false alarm); correctly indicating shoot in response to an armed target earned 10 points (hit), but indicating don’t shoot earned a penalty of 40 points (miss). Failure to respond within 630 ms of target onset resulted in a penalty of 10 points (timeout). Visual and auditory feedback, along with point totals, were presented after every trial. During the training phase, experts completed 16 practice trials and 200 training trials separated into two blocks. To match the level of exposure to the targets, undergraduates in the novice condition performed an observation task in which FPST images appeared under the same timing conditions as the expert condition. Novices were instructed to quickly press a button whenever the image included a person (i.e., no reference was made to shoot/don’t-shoot decisions). Police did not complete either training or observation task; they were considered experts by virtue of occupational training and on-the-job experience. During the test phase, all three groups completed 16 practice trials and 100 test trials of the FPST.

To obtain stable SDT estimates, we used a short response window (630 ms). This window typically induces a greater number of errors, which improves SDT estimates. However, a side effect of reducing the time window is that it reduces variance in response latencies, restricting our ability to detect effects in reaction time (see Correll et al., 2002). Past work indicates that training and expertise only affect the magnitude of bias in error rates, not latencies (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). Thus, the present studies all impose a 630 ms time window, and all analyses focus on SDT estimates.2

**Newspaper articles.** Adapted from Correll, Park, Judd, and Wittenbrink (2007), the articles described a string of extremely violent armed robberies. In the Black-criminal
condition, the article alleged that a pair of Black males between the ages of 30 and 35 committed the robberies. In the White-criminal condition, the content of the article was identical except the suspects were described as White males. Each story included police sketches of the ostensible suspects.

Procedure. Police officers were greeted by a male or female experimenter and seated at one of six workstations equipped with a laptop computer, a button box, and headphones. Undergraduates, in groups of one to three, were greeted by one of two female experimenters, and seated in cubicles equipped with a desktop computer and headphones. Experimenters outlined the study as an investigation of “responses and memory.”

Undergraduates either completed an initial 200-trial round of the FPST (experts) or simply observed the images (novices), ostensibly to familiarize them with the experimental stimuli. Police officers had no exposure to the FPST stimuli. All participants were then randomly assigned to either the White-criminal or Black-criminal condition and were given 5 min to study the corresponding article. Participants were told that there would be a memory test concerning the events and images described in the article. After reading the article, all participants completed the FPST. Following the test phase, participants were given 5 min to recall as many details as possible about the article.7 Last, participants completed demographic and individual difference measures.8 Participants were then debriefed and thanked for their time.

Results and Discussion

Participants responded incorrectly on 12.61% of the trials and timed out on another 13.28% of the trials. Correct and incorrect responses (excluding timeouts) were used to conduct signal-detection analysis. Applied to the FPST, SDT (D. M. Green & Swets, 1966) assumes that armed and unarmed targets vary along some dimension relevant to the decision at hand (e.g., the extent to which targets appeared threatening). SDT estimates participants’ ability to discriminate between targets with guns and without guns (d’) and the point on the decision-relevant dimension at which participants decide that a target is threatening enough to warrant shooting (c). Higher values of d’ indicate greater sensitivity. A criterion of zero indicates no tendency to favor either a shoot response or a don’t-shoot response. Deviations from zero in the positive direction indicate a conservative tendency to favor the don’t-shoot response, and deviations in the negative direction indicate a trigger-happy tendency to favor the shoot response.

The primary focus of the present studies concerns the effect of our manipulations on racial bias in the criterion to shoot. Previous work suggests that target race does not affect participants’ ability to differentiate between armed and unarmed targets (Correll et al., 2002); thus, we did not have strong hypotheses about the effect of the manipulations on sensitivity. However, we did predict that trained individuals would show greater sensitivity, on average, than untrained individuals as past studies have found that officer and expert samples show greater sensitivity than community members and novices (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007).

For each participant, we calculated c and d’ once for White targets and once for Black targets.3 We then computed the difference in the estimates between White and Black targets, and the average values of both c and d’ (i.e., across White and Black targets). The indices were submitted to separate between-subjects ANOVAs, testing two orthogonal contrasts for the effects of Training Condition. The first of these contrasts directly tests Hypothesis 1, comparing novices to all trained participants (novice = −2, expert = 1, police = 1). The second contrast compared expert undergraduates to police officers (novice = 0, expert = −1, police = 1), allowing us to determine whether the nature of the training matters. In addition, we examined the interactions between Article Condition and each Training Condition contrast. As in previous studies (Correll, Park, Judd, & Wittenbrink, 2007), we expected participants to show greater racial bias after reading about Black criminals than White criminals. But our primary hypothesis involved the capacity for training to attenuate this effect. Thus, we were particularly interested in the interaction between Article Condition and the first Training Condition contrast (which compares all trained participants with novices).

Racial Bias in Criteria. To quantify the magnitude of racial bias, we computed the difference between the criterion for White and Black targets (c White − c Black; see Table 1). Higher numbers indicate greater bias (i.e., a more conservative criterion for Whites, and/or a lower, more trigger-happy criterion for Blacks). On average and controlling for condition, participants demonstrated significant racial bias (M = 0.093, SD = 0.298), F(1, 107) = 10.772, p = .001. There were no main effects of Training or Article, though there was a weak trend for participants to show more bias after reading about Black criminals (M = 0.126, SD = 0.297) than White criminals (M = 0.059, SD = 0.298), F(1, 107) = 2.036, p = .157.

As predicted, we observed an interaction between Article Condition and the Training Condition contrast comparing novices to all trained participants. The articles had greater impact on novices than on trained individuals, F(1, 107) = 4.731, p = .032 (see Figure 1). Novices showed greater bias after reading about Black than White criminals, t(107) = 2.507, p = .014. But the articles had no significant impact on racial bias for either experts, t(107) = −0.365, p = .716, or police officers, t(107) = 0.218, p = .828. The interaction between Article Condition and the second Training Condition contrast (comparing experts with officers) was not significant, F(1, 107) = 0.176, p = .676, offering no evidence that the articles affected the two groups differently.6 Pairwise analyses indicate that the articles had greater impact on novices than on experts, F(1, 107) = 4.194,
Sim et al.

In sum, novices were dramatically affected by manipulations of stereotype accessibility, whereas experts and officers were not. These data provide fairly clear support for Hypothesis 1, suggesting that training can reduce the effect of accessible stereotypes.

In addition to the test of our primary hypothesis, we conducted a number of ancillary analyses to more fully explore the data. Though not central to our predictions, these tests are briefly described below.

**Mean Criterion.** Analysis of the mean criterion (the average of Black and White targets) revealed that across all conditions and regardless of target race, participants demonstrated a negligible tendency in favor of the shoot response \((M = -0.033, SD = 0.259), F(1, 107) = 2.199, p = .141\). Novices set more lenient criteria than trained individuals, indicating a greater overall willingness to shoot, \(F(1, 107) = 4.774, p = .031\). Expert undergraduates and police officers did not differ from each other, \(F(1, 107) = 0.133, p = .716\).

**Racial Bias in Sensitivity.** We did not predict that target race would affect sensitivity. However, on average and controlling for condition, participants demonstrated greater sensitivity to Black than White targets \((M = -0.163, SD = 0.596), F(1, 107) = 13.461, p < .001\). This tendency was somewhat weaker among participants who had read about Black criminals \((M = -0.266, SD = 0.568)\) than White criminals \((M = -0.062, SD = 0.610), F(1, 107) = 3.342, p = .07\).

---

Table 1. Decision Criterion and Sensitivity Means and Standard Deviations for Novice and Expert Participants, as well as for Police Officers Who Read About Either a Black or White Criminal, Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Novice Black criminal</th>
<th>Novice White criminal</th>
<th>Expert Black criminal</th>
<th>Expert White criminal</th>
<th>Police Black criminal</th>
<th>Police White criminal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criterion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.014</td>
<td>0.297</td>
<td>-0.152</td>
<td>0.371</td>
<td>0.084</td>
<td>0.291</td>
</tr>
<tr>
<td>Black</td>
<td>-0.207</td>
<td>0.212</td>
<td>-0.115</td>
<td>0.295</td>
<td>-0.033</td>
<td>0.222</td>
</tr>
<tr>
<td>Difference</td>
<td>0.220</td>
<td>0.234</td>
<td>-0.037</td>
<td>0.290</td>
<td>0.117</td>
<td>0.250</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2.335</td>
<td>0.646</td>
<td>1.838</td>
<td>0.744</td>
<td>2.176</td>
<td>0.673</td>
</tr>
<tr>
<td>Black</td>
<td>2.498</td>
<td>0.778</td>
<td>2.302</td>
<td>0.803</td>
<td>2.750</td>
<td>0.685</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.164</td>
<td>0.495</td>
<td>-0.464</td>
<td>0.526</td>
<td>-0.574</td>
<td>0.592</td>
</tr>
</tbody>
</table>

**Figure 1.** Racial bias in criteria (White–Black) as a function of Training Condition and Article Condition in Experiment 1

Note: Error bars represent standard errors.

\(*p < .05\).
Furthermore, Article Condition affected novices differently than trained individuals, $F(1, 107) = -10.214$, $p = .002$, and experts differently than officers, $F(1, 107) = 3.378$, $p = .069$. These unanticipated effects seem to be driven almost entirely by the performance of two particular groups who showed a surprising effect of target race: the novice undergraduates in the White-criminal condition, $t(107) = -3.224$, $p = .002$, and the expert undergraduates in the Black-criminal condition, $t(107) = -3.983$, $p < .001$. Race did not affect sensitivity for any other condition, $F$s $< 1.563$. In the absence of an a priori prediction or obvious theoretical meaning, and because the corresponding effect does not emerge in the subsequent studies, we hesitate to interpret this result.

**Mean Sensitivity.** Past research has found that training and expertise improve overall accuracy or sensitivity to the presence of a weapon (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). Somewhat surprisingly then, sensitivity did not vary by condition in these data, $F$s $< 2.359$.

**Experiment 2a**

Experiment 1 showed that lab-based training and real-world experience could reduce the impact of accessible stereotypes. Training and expertise may thus enable participants to attend to other cues and/or to override stereotypes (even when they are made salient by a newspaper article). The second set of experiments (Experiments 2a and 2b) probed a boundary condition for these beneficial effects. If practice alone is sufficient to explain the reductions in bias observed in Experiment 1, then training on the FPST should always reduce bias. However, it is possible (and, indeed, probable) that the training environment will moderate its efficacy. For example, if participants receive training in a task environment that reinforces cultural stereotypes—say, responding to a large number of armed Black targets and very few armed White targets—participants might come to expect that Blacks pose a greater threat than Whites.

Plant, Peruche, and Butz (2005) conducted two FPST studies that bear on this question. In one study (Study 1), participants were repeatedly exposed to Black and White targets, but race was unrelated to the presence or absence of a gun. At the beginning of the task, participants showed bias, shooting more Black than White targets. On the later trials of the task, the participants responded with less and less bias. These participants seemingly learned to ignore race as the study progressed. In a separate study (Study 3), the researchers presented participants with a stimulus set that reinforced the cultural stereotype. Unlike in Study 1, the race of the target was related to the presence/absence of a weapon, such that there were a disproportionately high number of Black-gun targets and White-no-gun targets. Here, the experimenters found no benefit of practice with the task (when comparing participants’ responses on early vs. late trials). When the stimuli reinforced the Black-danger stereotype, racial bias was not eliminated. In a similar fashion, Correll, Park, Judd, and Wittenbrink (2007) showed that stereotypic covariation in the stimulus set of the FPST increased bias in reaction times. In particular, a SC pattern of covariation, relative to a SI pattern of covariation, intensified the association between Blacks and danger and the magnitude of racial bias exhibited in reaction times (Correll, Park, Judd, & Wittenbrink, 2007, Studies 2 and 3).

The present experiments extended past work in two critical ways. First, unlike Plant and colleagues (2005), we experimentally manipulated the covariation between race and weapon within the same experimental context, allowing us to test whether covariation in the training stimuli is causally related to performance differences. Second, unlike Correll, Park, Judd, and Wittenbrink (2007), we focused primarily on bias in error rates rather than reaction time. As discussed, training and experience with the FPST seem to reduce bias in errors, but training has little to no effect on bias in reaction time. Accordingly, if we hope to understand how the task environment moderates the effect of training, we must examine error rates.

We hypothesized that when race is unrelated to the presence/absence of a weapon, training may reduce racial bias as participants learn to focus on diagnostic information (i.e., the object). When race covaries meaningfully with threat, however, “training” may teach people to rely on race to facilitate weapon detection. This was the crux of our second hypothesis. In Experiment 2a, we tested this hypothesis experimentally with students in the laboratory; in Experiment 2b, we examined this hypothesis among police officers, using correlational methods to look at differences in the nature of their on-the-job experiences.

**Method**

Participants were randomly assigned to a training task that either reinforced the Black-danger stereotype (i.e., where Black targets were disproportionately likely to be armed, and White targets were disproportionately likely to be unarmed), or undermined the stereotype (i.e., where Black targets were disproportionately likely to be unarmed and White targets were disproportionately likely to be armed), or was effectively race neutral. To the extent that participants are sensitive to the contingency between target race and object type, reductions in bias should only occur when people practice responding in a race-neutral or in a SI training context. However, participants who never learn to overcome the Black-danger stereotype (i.e., those who learn to expect a SC pattern of covariation in the training phase) should exhibit greater bias in errors than those who receive practice on a SI pattern of covariation. Stated simply, when the environment reinforces (rather than undermines) the Black-danger association, training may not reduce racial bias in the decision to shoot.
Participants and Design. A total of 120 students were recruited from the Chicago area, primarily from college, university, and technical/trade school campuses (62 female, 58 male; 46 White, 18 Asian, 49 Black, 7 Latino/a; M age = 23.76 years) to participate in the study for US$12. As in Experiment 1, the reported results are based on non-Black participants. Thus, the final sample consisted of 71 participants.

The experiment involved a 3 (Training Condition: SC vs. no-covariation control vs. SI) × 2 (Target Race: Black vs. White) × 2 (Object Type: gun vs. no gun) mixed-model design, with repeated measures on the last two factors in the test phase.

FPST. During the training block, participants completed 16 practice trials and 200 training trials. The FPST was modified for each of the three training conditions. One third of the participants were assigned to a no-covariation condition. In this control condition, Black and White targets were equally likely to be armed or unarmed. In other words, the trials were evenly divided between the four target types (armed Black, armed White, unarmed Black, unarmed White). In the remaining two training conditions, participants were exposed to either a SC or a SI pattern of covariation. In the SC training condition, participants were assigned to a no-covariation condition. In this control condition, Black and White targets were equally likely to be armed or unarmed. In other words, the trials were evenly divided between the four target types (armed Black, armed White, unarmed Black, unarmed White). In the remaining two training conditions, participants were exposed to either a SC or a SI pattern of covariation. In the SC training condition, Black targets in the FPST were particularly likely to be armed and White targets were particularly likely to be unarmed. In the SI training condition, Black targets in the FPST were particularly likely to be armed and White targets were particularly likely to be unarmed (see Correll, Park, Judd, & Wittenbrink, 2007, for details). During the test block, all participants completed 100 test trials of the standard FPST. It is critical to note that, regardless of the participant’s training condition, this test phase involved no covariation between target race and object type. The point values and the response window were identical to Experiment 1.

Results and Discussion

Overall, participants responded incorrectly on 16.48% of the trials and timed out on another 12.56% of the trials. Correct and incorrect responses (excluding timeouts) were used to conduct signal-detection analysis. As in Experiment 1, we calculated c′ and d′ separately for White targets and Black targets, then we computed the averages and difference scores (White − Black) for both c′ and d′. The indices were submitted to separate between-subjects ANOVAs, specifying two orthogonal contrasts for the effects of Training Condition. In the present study, we were concerned with the covariation between race and threat, which increased linearly from the SI condition (which undermines stereotypes) to the no-covariation control condition (which is neutral with respect to stereotypes) to the SC condition (which reinforces stereotypes). Our second hypothesis was tested by the linear effect of Training Condition (SI = −1, control = 0, SC = 1), which determines whether racial bias increases as a function of the covariation between race and threat during the training phase. The quadratic or residual contrast, which compared the no-covariation control condition to the two experimental covariation conditions (SI = −1, control = 2, SC = −1), completed the set of orthogonal contrasts but had no particular theoretical relevance to our hypothesis.

Racial Bias in Criteria. Again, we computed the SDT index of racial bias (c′_{White} − c′_{Black}); see Table 2 for complete means and standard deviations). On average, controlling for condition, participants demonstrated bias (M = 0.160, SD = 0.414), F(1, 68) = 11.102, p = .001, setting a higher (more conservative) criterion for Whites and a lower (more lenient or trigger-happy) criterion for Blacks.

Critically, the linear contrast for Training Condition significantly predicted bias, F(1, 68) = 4.289, p = .042 (see Figure 2). As anticipated, participants exposed to SC training showed more bias than participants exposed to the SI training. The quadratic or residual contrast was not significant, suggesting that the no-covariation control condition fell between the two other training conditions and did not differ statistically from their average, F(1, 68) < .001, p = .987.

The data suggest that racial bias in the test phase varied as a function of Training Condition. Participants in the SC condition showed significant bias, setting a more lenient criterion for Black than White targets, t(68) = 3.301, p = .002.

Table 2. Decision Criterion and Sensitivity Means and Standard Deviations for SI, No-Covariation Control, and SC Training Conditions, Experiment 2a

<table>
<thead>
<tr>
<th></th>
<th>SI</th>
<th>Control</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>−0.054</td>
<td>0.099</td>
<td>0.134</td>
</tr>
<tr>
<td>Black</td>
<td>−0.089</td>
<td>−0.063</td>
<td>−0.153</td>
</tr>
<tr>
<td>Difference</td>
<td>0.035</td>
<td>0.162</td>
<td>0.287</td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1.772</td>
<td>1.758</td>
<td>1.829</td>
</tr>
<tr>
<td>Black</td>
<td>2.309</td>
<td>2.119</td>
<td>1.784</td>
</tr>
<tr>
<td>Difference</td>
<td>0.538</td>
<td>−0.361</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Note: SI = stereotype-incongruent; SC = stereotype-congruent.

Downloaded from psp.sagepub.com at UNIV OF COLORADO LIBRARIES on June 18, 2013
Somewhat surprisingly, participants in the control condition also displayed bias, $t(68) = 2.035, p = .046$. Participants in the SI condition, however, showed no evidence of racial bias, $t(68) = 0.414, p = .680$.

In support of Hypothesis 2, participants in the SC condition acquired experience with the FPST, but their training did not eliminate subsequent bias. Presumably, mere practice with shoot/don’t-shoot decisions in response to a series of SC targets did not eliminate bias because the racial information in the training phase served as a valid predictor of the presence (or absence) of a gun. Indeed, racial bias in the test phase was only eliminated when participants practiced with a series of SI targets during training.

As in Study 1, we conducted a series of ancillary analyses to more fully explore the dataset.

**Mean Criterion.** Analysis on the mean criterion level revealed that across all conditions and regardless of target race, participants demonstrated a negligible tendency to favor the shoot response ($M = -0.020, SD = 0.236$), $F(1, 68) = 0.557, p = .458$. Training Condition had no effect on overall willingness to shoot: Neither the linear contrast, $F(1, 68) = 0.780, p = .380$, nor the quadratic contrast, $F(1, 68) = 1.000, p = .321$, affected this tendency.

**Racial Bias in Sensitivity.** Overall, participants showed greater sensitivity to Black than White targets ($M = -0.292, SD = 0.727$), $F(1, 68) = 11.765, p < .001$. This difference in sensitivity was moderated by the linear contrast, $F(1, 68) = 7.879, p = .007$, the quadratic contrast was not significant, $F(1, 68) = 0.450, p = .505$. The difference in sensitivity was larger in the SI than in the SC condition. This effect was largely driven by responses to the Black targets, which depended on the linear contrast, $F(1, 68) = 4.268, p = .043$, again, the quadratic was not significant, $F(1, 68) = 0.120, p = .730$. Sensitivity to Black targets was lower in the SC condition than the SI condition, where participants were presumably forced to overcome cultural stereotypes. Sensitivity to White targets was unrelated to condition, $F(< .056)$.

**Mean Sensitivity.** Analysis of the mean sensitivity level (the average of Black and White targets) revealed no effects of either the linear, $F(1, 68) = 1.075, p = .303$, or the quadratic contrast, $F(1, 68) = 0.007, p = .935$.

In essence, SI participants, who were exposed to counterstereotypic stimuli during a training phase, showed higher sensitivity to Black targets, but this sensitivity decreased as the training stimuli became more congruent with the cultural stereotypes (i.e., in the SC condition). It may be the case that exposure to the SI stimuli caused participants to focus extra attention on Black targets.

Together, the SDT analyses indicate that mere repetition cannot explain the benefits of training. Performance on the standard FPST appears to depend critically on the content presented during the training phase. This is consistent with the possibility that training is most effective at reducing bias when the trainee must frequently ignore or override relevant stereotypes to execute the correct response. If stereotypes continually provide a valuable heuristic in the training environment, then even trained individuals may use them. As suggested by Hypothesis 2, this latter form of “training” may exacerbate, not attenuate, bias.
Experiment 2b

The previous study demonstrated that, for trained participants, racial bias in the FPST depends on the nature of stimuli presented during training. Through practice, participants seem to learn about their environment and presumably form expectations (either implicit or explicit) about what cues they ought to use or ignore. Participants exposed to environments in which race is unrelated to threat may come to disregard race; those exposed to environments in which Blacks reliably pose a greater threat may learn to use race as a diagnostic cue (Correll, Park, Judd, & Wittenbrink, 2007; Plant et al., 2005). What are the practical implications of this finding? Like anybody else, police officers presumably learn from experience, so even highly trained officers might show differences in performance to the extent that their on-the-job experiences differentially reinforce racial stereotypes. Over the past 8 years, we have tested almost a thousand officers from all over the country. The vast majority of these officers are “beat cops” who patrol a particular region of their city, interacting with witnesses or victims of crime, performing traffic stops, among other duties. These officers have extensive contact with a variety of people, most of whom are (for lack of a better term) good guys (National Institute of Justice, 2010). In the course of our research, we have also had the opportunity to test a small subset of officers (n = 22) whose jobs require them to specifically target and interact with violent criminals. Their responsibilities include preventing and reducing gang violence, which often involves interacting with and arresting violent gang members. Members of these special units (SUs), which deal primarily with gangs or street crime, may have much greater contact with bad guys, and to the extent that these units deal primarily with Blacks and other minorities, the officers’ on-the-job experiences may tend to reinforce cultural stereotypes. These SU officers may thus experience a much more dramatic and consequential real-world version of the kind of stereotype-reinforcing environment explored in Experiment 2a.

In Experiment 2b, we examined these 22 SU officers, whose data have not been reported elsewhere. By way of comparison, we also present data from patrol officers and untrained members of the community, which were previously reported by Correll, Park, Judd, Wittenbrink, Sadler, et al. (2007, Study 2). Although the experimental procedures and materials were identical across all three samples, it is important to note that the SU officer and patrol officer/community data were collected at different times and locations, so comparisons should be interpreted with caution.

As discussed, trained participants and officers tend to show less racial bias on the criterion measure than novices (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). But, in line with Experiment 2a, the capacity of an officer’s training and experience to reduce bias may depend heavily on the degree to which the environment reinforces the stereotype that Blacks are dangerous. For officers who routinely deal with (stereotype-congruent) minority gang members, race may be taken as a diagnostic cue, leading to more extreme bias in the FPST.

Method

Participants and Design. Twenty-two SU officers from gang and street-crime units (1 female, 21 male; 12 White, 7 Black, 2 Latino/a, 1 Native American/Pacific Islander; M age = 39.4 years) were tested. These data were compared with data from 31 patrol officers (3 female, 26 male, 2 missing gender; 16 White, 6 Black, 4 Latina/o, 3 Other; 2 missing ethnicity; M age = 35.6 years) and 45 community members (20 female, 23 male, 2 missing gender; 14 White, 18 Black, 10 Latina/o, 3 Other; M age = 36.8 years) from Denver, Colorado (data originally reported in Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). Community members were recruited from the Department of Motor Vehicles and were compensated US$20. Officers completed the study during off-duty hours for US$50.

All participants completed the standard FPST. The experiment involved a 3 (Sample: SU officer vs. patrol officer vs. community member) × 2 (Target Race: Black vs. White) × 2 (Object Type: gun vs. no gun) mixed-model design, with repeated measures on the last two factors.

Procedure. SU officers were greeted by a male or female experimenter and seated at one of six separate workstations equipped with a laptop computer, a button box, and head-phones. Similar procedures were used for patrol officers and community members (for details, see Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007, Study 2). All three groups of participants completed 16 practice trials and 100 test trials of the standard FPST (identical to the test phase of Experiments 1 and 2a). Following the FPST, participants completed demographic and individual difference measures. Participants were then debriefed and thanked for their time.

Results and Discussion

Overall, participants responded incorrectly on 16.15% of the trials and timed out on another 17.94% of the trials. Correct and incorrect responses (excluding timeouts) were used to conduct signal-detection analysis. Again, we calculated c and d’ for White and Black targets, and computed the race difference and the average. Based on our theoretical predictions, our analyses focused on pairwise comparisons between the samples.

Racial Bias in Criteria. We computed the SDT index of racial bias for each sample (see Table 3). On average, participants showed bias (M = .076, SD = 0.364), F(1, 95) = 4.347, p = .040. As reported in Correll, Park, Judd, Wittenbrink, Sadler, et al. (2007), patrol officers displayed less
bias than community members, \(F(1, 95) = 4.068, p = .046\).

The critical comparisons, here, involved comparisons between the SU officers and the two other samples. If SU officers’ on-the-job experience with gang members and street crime create a “stereotype-congruent” environment that allows them to rely more heavily on heuristics (rather than exercising constant control), they may show more bias than patrol officers (whose day-to-day experience may be less congruent with cultural stereotypes). Indeed, SU officers showed significantly more bias than patrol officers, \(F(1, 95) = 4.735, p = .032\) (see Figure 3). In fact, SU officers showed levels of bias that were comparable to (and even nonsignificantly greater than) untrained community members, \(F(1, 95) = 0.271, p = .603\). This pattern offers tentative real-world support for Hypothesis 2, that when experience reinforces racial stereotypes, training and expertise may not reduce bias.

After testing our primary hypothesis, we again conducted a number of ancillary analyses.

**Mean Criterion.** Across all conditions and regardless of target race, participants favored the shoot response (\(M = -0.156, SD = 0.279\), \(F(1, 95) = 21.800, p < .001\)). Community members (\(M = -0.244, SD = 0.302\)) set more lenient criteria than both patrol officers (\(M = -0.098, SD = 0.248\), \(F(1, 95) = 5.406, p = .022\)), and SU officers (\(M = -0.058, SD = 0.222\), \(F(1, 95) = 7.007, p = .009\)), suggesting a greater overall willingness to shoot. Patrol officers did not differ from SU officers, \(F(1, 95) = 0.275, p = .602\). This finding corroborates previous research that has found that police officers set more conservative criteria overall than community members.

**Racial Bias in Sensitivity.** Overall, participants did not show differential sensitivity to Black or White targets (\(M = -0.057, SD = 0.750\), \(F(1, 95) = 1.166, p = .283\)). SU officers did not differ from community members, \(F(1, 95) = 1.153, p = .286\), or patrol officers, \(F(1, 95) = 0.213, p = .645\).

**Mean Sensitivity.** Community members (\(M = 1.431, SD = 0.836\)) were less sensitive than both patrol officers (\(M = 2.279, SD = 0.695\), \(F(1, 95) = 23.002, p < .001\)), and SU officers (\(M = 2.088, SD = 0.665\), \(F(1, 95) = 11.116, p = .001\)). Patrol officers did not differ from SU officers, \(F(1, 95) = 0.817, p = .368\). In line with past research (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007), the data suggest that police officers were more accurate than community members.

The analyses suggest that when one’s environment or professional context reinforces the utility of race information, even highly trained individuals show racial bias (as measured by criterion estimates). This occurs despite other observable benefits of training. SU officers outperformed community members on a number of measures, showing greater sensitivity to the presence of a weapon and setting more conservative criteria for the decision to shoot. However, like untrained community members, race clearly impacted the SU officers’ criteria: They were significantly more likely to shoot Black than White targets.

**General Discussion**

The fact that police shoot a disproportionate number of Black suspects has led some to conclude that the “police have one trigger finger for Whites, and another for Blacks” (Takagi, 1974). At the same time, when compared with lay participants, officers generally show less racial bias in laboratory-based shooter simulations, perhaps as a function of their training and expertise (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). The present research explored whether and how training and expertise might reduce bias, and when those beneficial effects are likely to fail.

Experiment 1 used different versions of a newspaper article to link Blacks or Whites to the concept of danger and crime. When novices read articles about Black criminals, they showed pronounced racial bias; when they read about White criminals, bias was eliminated. In essence, novices were highly sensitive to the manipulation of stereotype accessibility. In contrast, expert undergraduates (who practiced the FPST) and police officers were essentially unaffected by the newspaper manipulation. These data suggest that race-neutral training enables participants to either ignore or override salient racial information, and learn to focus on task-relevant cues.

Experiments 2a and 2b, however, suggest that the nature of training moderates participants’ ability to overcome or ignore stereotypes. When the training context (Experiment 2a) or the nature of on-the-job experiences (Experiment 2b) reinforced the association between Blacks and danger, training did not seem to attenuate bias. When race is diagnostic of danger (e.g., in certain neighborhoods; White, 2002), in a sense, when race becomes a task-relevant cue, even experts may rely on it to facilitate weapon detection.

### Table 3. Decision Criterion and Sensitivity Means and Standard Deviations for Community Members, “Beat” Patrol Officers, and SU Officers, Experiment 2b

<table>
<thead>
<tr>
<th></th>
<th>Community members</th>
<th>“Beat” officers</th>
<th>SU officers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M)</td>
<td>SD</td>
<td>(M)</td>
</tr>
<tr>
<td><strong>Criterion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>-0.185</td>
<td>0.386</td>
<td>-0.123</td>
</tr>
<tr>
<td>Black</td>
<td>-0.303</td>
<td>0.328</td>
<td>-0.073</td>
</tr>
<tr>
<td>Difference</td>
<td>0.118</td>
<td>0.385</td>
<td>-0.050</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1.474</td>
<td>1.026</td>
<td>2.170</td>
</tr>
<tr>
<td>Black</td>
<td>1.388</td>
<td>0.843</td>
<td>2.388</td>
</tr>
<tr>
<td>Difference</td>
<td>0.086</td>
<td>0.855</td>
<td>-0.218</td>
</tr>
</tbody>
</table>

Note: SU = special unit.
Possible Mechanisms Involved in Training

There are several possible mechanisms that could underlie the advantages of training. The first possibility is that with training, individuals learn to inhibit the activation of race. Stereotypes often come to mind without awareness or intention (Macrae, Milne, & Bodenhausen, 1994). While simplifying and streamlining the processing of SC targets (e.g., armed Blacks, unarmed Whites), stereotypes can impede the processing of SI targets (e.g., unarmed Blacks, armed Whites) and affect the construal of ambiguous information (Kunda & Sherman-Williams, 1993). When stereotypes are accessible, novices may be more likely to mistake innocuous objects for guns when those objects are paired with Black (rather than White) targets. Training or extensive practice may reduce the activation of racial concepts. Indeed, in one study, after completing a FPST in which race was nondiagnostic, participants were less likely to complete word fragments with race-related words (Plant et al., 2005).

Another possibility is that, with practice, individuals learn to automatically base their decisions to shoot on the object itself (i.e., presence or absence of a gun). With training, individuals may become so skilled at object-based judgments that the process becomes extremely rapid and efficient (Shiffrin & Schneider, 1977). Although these two explanations are reasonable accounts for some existing data, the idea that trained participants (or experienced officers) suppress race-based processing or process weapons “automatically” cannot easily account for the present data. If either possibility were a viable explanation, officers and trained undergraduates should demonstrate less racial bias in both their willingness to shoot and their reaction times. But even highly trained officers in our past studies (who show dramatic reductions in bias as assessed through SDT criteria) still show pronounced bias in their reaction times (Correll, Park, Judd, Wittenbrink, Sadler, et al., 2007). This simple fact strongly suggests that the officers are both attending to racial cues and activating race-based stereotypes, leading them to see Blacks as potential threats and facilitating the decision to shoot.

If training and expertise do not eliminate either attention to race or task-relevant racial stereotypes, how do these experiences attenuate bias? We tentatively propose that the critical process involves expertise-based increases in cognitive control. This third possible mechanism nicely accounts for our data. Even though experts may process racial cues and access racial stereotypes—indeed, even when those stereotypes are rendered highly accessible through experimental manipulations as in Experiment 1—control may allow experts to execute an unbiased response. If extensive practice and training enhances control, manipulations that undermine cognitive control should compromise the value of training. Indeed, a recent investigation provides evidence that cognitive load (i.e., depriving participants of cognitive resources) compromises the performance of experts, leading to an increase in racial bias (Correll, Wittenbrink, Axt, Goyle, & Miyake, 2012).

Practice-related changes can be examined on the behavioral, cognitive, and neural level (Kelly & Garavan, 2005). Cognitive control may enable individuals to either (a) more effectively extract diagnostic information from the stimulus (e.g., increased visual attentional capacity; C. S. Green & Bavelier, 2003) or (b) more heavily weight diagnostic information in executing the ultimate response. In the
FPST, the decision to shoot is typically affected by both the presence of a gun (the relevant cue) and the race of the target (a robust, but irrelevant cue). When individuals perceive that the object is inconsistent with the stereotype brought to mind by the race of the target (e.g., unarmed Black), the conflict in potential response (shoot vs. don’t shoot) may stimulate the anterior cingulate cortex (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999) and trigger executive control activity in the prefrontal cortex (Miller & Cohen, 2001). Indeed, event-related brain potentials reveal that individuals who are better at detecting response conflict are more able to exert cognitive control over their decisions (Amodio et al., 2004).

However, according to this account, training and expertise should only minimize bias when they focus on nonracial information—when training requires the participant to respond in a counterstereotypic fashion. If the training environment actually promotes the diagnostic utility of racial cues (rather than the utility of increased control), training and expertise should have little to no beneficial effect against the validity of stereotype-congruent heuristic responses—exactly the pattern of behavior observed in Experiments 2a and 2b.

Limitations and Future Directions

The FPST is an impoverished simulation of an actual police encounter. The images are static, bystanders are deliberately precluded, trials occur in rapid succession, and responses are made by pressing buttons (not by pulling triggers). The task assesses bias in a split-second reaction that would ordinarily represent the final stage in police decision making. Perhaps most critically, the FPST takes place in the safety of a lab, and failure in the simulation poses no real threat.

In real-world confrontations, officers who are contemplating lethal force necessarily believe that the suspect might pose a lethal threat. This is an extraordinarily high-stress situation (Anshel, 2000). The constraints of time, cognitive resources, and ambiguity surrounding the level of danger present (Fridell & Binder, 1992) may diminish attentional capacity and escalate reliance on peripheral cues (Lambert et al., 2003; but see Loftus, Loftus, & Messo, 1987). Officer fatigue and sleep deprivation (e.g., from shift irregularities, overtime; Vila, Kenney, Morrison, & Reuland, 2000) may also increase reliance on stereotypic associations (Bodenhausen, 1990; Govorun & Payne, 2006; Ma et al., 2012). These factors limit the extent to which officers perceive and process information that might be important to the decision to shoot. Furthermore, if training-based reductions in bias are mediated by effortful cognitive control, it is possible that—when an officer is scared, stressed, or tired—his or her ability to recruit that control will be compromised (Beilock, 2010). Accordingly, a critical question for this research is to explore the effects of training and expertise on performance in nonoptimal test situations.

Conclusion

In any case of police-involved shooting, a multitude of variables are at play. By including training as a variable in empirical investigations, we can begin to hone in on the interactions that determine the decision to shoot, and with time, we may be able to isolate the mechanisms that enhance officer and community safety.

Acknowledgment

We wish to thank the members of the Stereotyping & Prejudice Research Laboratory at the University of Chicago for their helpful comments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Primary support for this work was provided by National Science Foundation Continuing Grant 0642580.

Notes

1. Past research suggests that Black participants may not be immune to the influence of prevalent cultural stereotypes—Black and White participants alike show racial bias in the decision to shoot (Correll, Park, Judd, & Wittenbrink, 2002); however, it is unclear how Black participants will react to a deliberate rendering of negative ingroup stereotypes. Where appropriate, we report analyses including all participants in the notes.

2. In all three studies, we computed bias in reaction time by averaging the log-transformed latencies from correct responses in the first-person-shooter task (FPST; see Correll et al., 2002). Although the means were in the expected direction (Experiment 1: $M = 0.011$, $SD = 0.026$; Experiment 2a: $M = 0.010$, $SD = 0.053$; Experiment 2b: $M = 0.005$, $SD = 0.033$), bias was not significant on average, except in Experiment 1: $F(1, 107) = 19.342$, $p < .001$; Experiment 2a and Experiment 2b: $F < 2.190$, nor was it moderated by condition or sample ($F < 1.812$).

3. Prior to the memory test, undergraduates were presented with 80 recognition trials to determine whether experts and novices were similarly acquainted with the targets presented. Each trial consisted of a target embedded in a particular background. Half of the presented images had in fact been seen previously; half had not. Participants pressed the “Y” key if they believed they had encountered the exact same image before and the “N” key if they did not recognize the image. There was no time limit for the recognition task. The number of correctly identified trials did not differ by experimental condition, suggesting that novices and experts had similar levels of familiarity with the targets ($F < 1.126$).
4. In all three studies, the individual difference measures do not shed light on our research questions so they will not be discussed further.

5. \( c = -0.5 \times (zFA + zH); d' = zH - zFA \), where \( FA \) is the proportion of false alarms (relative to correct rejections) and \( H \) represents the proportion of hits (relative to misses). The \( z \) operator is the translation of these proportions to \( z \) scores. Both \( FA \) and \( H \) were assigned a minimum value of 1 / 2n (where \( n \) is the total number of non-gun and gun trials, respectively), and a maximum of 1 – (1 / 2n) to eliminate infinite \( z \) scores.

6. The results were similar when Black participants were included in the analysis. The ANOVA revealed that on average and controlling for condition, participants demonstrated significant racial bias (\( M = 0.104, SD = 0.293 \)), \( F(1, 121) = 16.225, p < .001 \). There were no main effects of Training or Article, though once again there was a weak trend for participants to show more bias after reading about Black (\( M = 0.139, SD = 0.297 \)) than White criminals (\( M = 0.070, SD = 0.288 \)), \( F(1, 121) = 2.256, p = .136 \). Consistent with predictions, an Article \( \times \) Training interaction emerged such that the article had marginally greater impact on novices than on trained individuals, \( F(1, 121) = 3.316, p = .071 \). Novices showed greater bias after reading about Black than White criminals, \( t(121) = 2.249, p = .026 \). But the articles had no impact on racial bias for either experts, \( t(121) = -0.163, p = .871 \), or police officers, \( t(121) = 0.407, p = .685 \).

7. The results were similar when Black participants were included in the analysis. The ANOVA revealed that participants generally demonstrated significant racial bias (\( M = 0.169, SD = 0.398 \)), \( F(1, 117) = 21.921, p < .001 \). Again, the magnitude of bias in the test phase varied as a function of the covariation training manipulation during the training phase. As expected, participants exposed to stereotype-congruent (SC) training showed more bias than participants exposed to stereotype-incongruent (SI) training, \( F(1, 117) = 4.783, p = .031 \), while the control condition fell between the two other training conditions and did not differ from their average, \( F(1, 117) = 1.124, p = .291 \). Participants in the SC condition showed significant racial bias in the test phase suggesting that they set a more lenient criterion for Black than White targets, \( t(117) = 4.742, p < .001 \). Consistent with the predictions, participants in the control condition displayed only a marginal level of bias, \( t(117) = 1.838, p = .069 \), and participants in the SI condition showed no racial bias, \( t(117) = 1.57, p = .119 \).

References


