

## Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot

Joshua Correll  
University of Chicago

Bernadette Park and Charles M. Judd  
University of Colorado at Boulder

Bernd Wittenbrink  
University of Chicago

Melody S. Sadler  
University of Colorado at Boulder

Tracie Keese  
University of Denver

Police officers were compared with community members in terms of the speed and accuracy with which they made simulated decisions to shoot (or not shoot) Black and White targets. Both samples exhibited robust racial bias in response speed. Officers outperformed community members on a number of measures, including overall speed and accuracy. Moreover, although community respondents set the decision criterion lower for Black targets than for White targets (indicating bias), police officers did not. The authors suggest that training may not affect the speed with which stereotype-incongruent targets are processed but that it does affect the ultimate decision (particularly the placement of the decision criterion). Findings from a study in which a college sample received training support this conclusion.

*Keywords:* police, race, bias, weapon, training

Inspired in part by high-profile police shootings of unarmed Black men, a flurry of social psychological research has attempted to assess the influence of a suspect's race on the use of force, specifically in terms of the decision to shoot (Correll, Park, Judd,

& Wittenbrink, 2002; Greenwald, Oakes, & Hoffman, 2003; Payne, 2001). Although social psychologists have only recently addressed this question, the impact of suspect ethnicity on police shootings has long been the focus of researchers in other fields of study, particularly sociology, political science, and law enforcement. Investigators have consistently found evidence that police use greater force, including lethal force, with minority suspects than with White suspects (e.g., Inn, Wheeler, & Sparling, 1977; Smith, 2004; see Geller, 1982, for a review). Data from the Department of Justice (2001), itself, indicate that Black suspects are approximately five times more likely than White suspects, per capita, to die at the hands of a police officer.

One of the most detrimental consequences of police shootings is the upheaval they can provoke. Shootings of a minority suspect may engender a sense of mistrust and victimization among community members and give rise to conflict between the community and police. Weitzer and Tuch (2004) present evidence that members of ethnic minorities often feel that they are mistreated by the police, even after statistically controlling for factors like personal and vicarious experiences with the law, exposure to the media, and neighborhood disadvantage (see also Sunshine & Tyler, 2003). The implication is that the police are racist and that officers use excessive force with minority suspects. In response, Black people may engage in more belligerent behavior, including "talking back" to police officers, and—in a vicious cycle—this belligerence may prompt more severe use of force by police (Reisig, McCluskey,

---

Joshua Correll, Department of Psychology, University of Chicago; Bernadette Park, Charles M. Judd, and Melody S. Sadler, Department of Psychology, University of Colorado at Boulder; Bernd Wittenbrink, Graduate School of Business, University of Chicago; Tracie Keese, University of Denver.

Primary support for this work was provided by a grant from the Russell Sage Foundation. Support for this work also came from National Institute of Mental Health Grant F31-MH069017 to Joshua Correll and National Institute of Mental Health Grant R01-45049 to Bernadette Park and Charles M. Judd.

In the interest of disclosure, we note that Tracie Keese also serves as a commander in the Denver Police Department. We thank Chief Gerald Whitman, the Denver Police Department, Calibre Press, the Denver Department of Motor Vehicles, and (especially) the many officers of the Denver Police Department and police departments around the country for their assistance, patience, and participation. We also thank Alinne Barrera, Heather Coulter, and David M. Deffenbacher for their invaluable assistance with this research and Myron Rothbart for his many helpful comments.

Correspondence concerning this article should be addressed to Joshua Correll, Department of Psychology, University of Chicago, Chicago, IL 60637. E-mail: jcorrell@uchicago.edu

Mastrofski, & Terrill, 2004). It is equally important to note that, as a consequence of this tension, officers who see their job as protecting the community may feel, and to some extent may actually *be*, thwarted in their efforts to perform their duty.

Officer-involved shootings, then, can have severe consequences, not just for the officers and suspects involved, but for the community at large as well. It is of paramount importance to understand and explain why minority suspects are disproportionately likely to be shot. The sociological literature offers a number of explanations. Some research suggests that bias in police shootings stems, at least in part, from the officers' role as protectors of the privileged (predominantly White) classes over the less fortunate (predominantly minority) members of society (Sorenson, Marquart, & Brock, 1993). Others argue that the racial discrepancy in officer-involved shootings stems from differential minority involvement in criminal activity (Department of Justice, 2001; Inn et al., 1977) or from the fact that minorities are disproportionately likely to live and work in low-income, high-crime communities (Terrill & Reisig, 2003).

A primary strength of the sociological approach is that it examines police use of force directly and in its true context. These researchers study real locations and real officers, and their dependent variable is the number of suspects who are actually shot. They thus maintain the richness and complexity of the real world when analyzing relationships between officer-involved shootings and variables like race or community disadvantage. At the same time, the preexisting correlations among these variables confound efforts to assess their independent effects. For example, the relationship between the proportion of Black citizens in a community and perceptions of disorder (Sampson & Raudenbush, 2004) is inextricably tied to, and cannot be fully separated from, racial discrepancies in officer-involved shootings (Terrill & Reisig, 2003). For this reason, a social psychological analysis of the problem with experimental methods is useful not to replace but rather to supplement research of a more naturalistic sort.

Over the past several years, social psychological researchers have examined the effect of race on shoot/don't-shoot decisions using videogame-like simulations. In one paradigm, participants view a series of images (background scenes and people) and are instructed to respond to armed targets with a *shoot* response, and to unarmed targets with a *don't-shoot* response as quickly and as accurately as possible (Correll et al., 2002; Correll, Park, Judd, & Wittenbrink, 2007; Correll, Urland, & Ito, 2006). The results of some 20 studies with this task, with a variety of parameters and manipulations, consistently show racial bias in both the speed and accuracy with which such decisions can be made. Participants are faster and more accurate when shooting an armed Black man rather than an armed White man, and faster and more accurate when responding "don't shoot" to an unarmed White man rather than an unarmed Black man. The bulk of this research has been conducted with college students, but the effect has been replicated with community samples of both White and Black participants, and conceptually similar effects have been obtained by a number of other labs (Amodio et al., 2004; Greenwald et al., 2003; Payne, 2001; Payne, Lambert, & Jacoby, 2002; Plant, Peruche, & Butz, 2005). These findings, along with reports from sociological and related literatures, clearly indicate that race can play an important role in decisions about the danger or threat posed by a particular

person. But experimental data rarely speak directly to police behavior.

In our literature review, we discovered only two papers that examine officers in experimental studies of racial bias. Eberhardt, Goff, Purdie, and Davies (2004) found that priming the concept of crime served to orient attention to Black (more than White) faces. This pattern held for officers and civilians alike. Plant and Peruche (2005) examined training effects among officers on a task where images of White and Black men appeared with a gun or nongun object superimposed on the face. They found that officers showed racial bias in their errors during the first phase of the study (i.e., officers were more likely to mistakenly shoot Black targets who appeared with nongun objects, and to not shoot White targets who appeared with a gun in the first 80 trials of the task), but that bias fell to nonsignificant levels in the second phase (i.e., the last 80 trials of the task). These studies suggest that officers, like undergraduates, show racial biases in the processing of crime-related stimuli.

But there is reason to believe that police will differ from citizens in shoot/don't-shoot decisions. Most notably, officers receive extensive experience with firearms during their academy training (before they are sworn in) and throughout their careers. For example, the Denver Police Department requires that new recruits spend 72 hr in practical weapons training, and officers must recertify on a quarterly basis. At the firing range, officers and recruits make shoot/don't-shoot decisions for target silhouettes that appear suddenly, either armed or unarmed; in Firearms Training System simulators (Firearms Training Systems, Inc., Atlanta, GA), they respond to an interactive video simulation of a potentially hostile suspect; and in simulated searches, they confront live actors armed with weapons that fire painful but nonlethal ammunition (e.g., paintballs, Simunition, or Air Soft pellets).

An extensive body of research shows that training improves performance on tasks in which a peripheral cue interferes with a participant's response to a central or task-relevant cue. Through training, participants learn to ignore the irrelevant information and respond primarily on the basis of the central feature of the stimulus (e.g., MacLeod, 1998; MacLeod & Dunbar, 1988; Plant & Peruche, 2005). For example, in a Stroop (1935) task, participants classify the color in which a word is printed (e.g., red). Color is thus the central cue. This task becomes more difficult if the word (a peripheral cue) refers to a different color (e.g., the word "blue" printed in red). Initially, participants have difficulty with this task, responding slowly and inaccurately when the central and peripheral cues conflict. But with training, judgment improves. Responses occur more quickly and require less effort and less cognitive control. As a result, experts demonstrate reduced interference in both latencies and errors. Neuroimaging studies have even documented the shifting patterns of brain activity that correspond to the development of automatic task performance (Bush et al., 1998; Jansma, Ramsey, Slagter, & Kahn, 2001; for a review, see Kelly & Garavan, 2004). During initial performance on interference tasks, participants recruit brain regions related to conflict detection and response control (e.g., the anterior cingulate and medial prefrontal cortexes). With extensive practice, however, activation in these regions decreases, presumably because an automatic task requires less executive supervision.

But automatization may not characterize all learning on interference tasks. In some cases, training actually promotes *controlled*

*processing*. For example, when participants are continuously challenged by variable task requirements or increasing demands, practice can lead to more extensive recruitment of prefrontal brain regions (Olesen, Westerberg, & Klingberg, 2004; Weissman, Woldorff, Hazlett, & Mangun, 2002). Of particular relevance to shoot/don't-shoot decisions, this control involves the medial and middle frontal gyri—areas related to the detection and resolution of conflicting information and to the maintenance of goal-relevant representations. In some cases, then, training leads participants to work harder, in cognitive terms, as they learn to marshal the attention and control necessary for optimal performance.

When will training promote automaticity in a judgment task, and when will it promote control? A probable moderator is task complexity (Birnboim, 2003; Green & Bavelier, 2003). On tasks with simple stimuli (e.g., the words presented in a Stroop task), practice allows participants to streamline the judgment process, performing it easily and automatically. Only when the task is difficult (e.g., involving visually complex stimuli or ever-changing task requirements) does practice seem to promote control. As Birnboim (2003) wrote, “automatic processing relies on a reduction of stimulus information to its perceptual and motor features” (p. 29). When complexity renders this kind of reduction impossible, controlled processing may be required to “extract more meaningful information” (p. 29). Consistent with this argument, Green and Bavelier (2003) have shown that practice on a visually complex video game (i.e., *Medal of Honor*; Electronic Arts, Redwood City, CA) improves performance on attention-demanding tasks, but practice on a visually simple video game (i.e., *Tetris*; Electronorgtechnica, Moscow, Russia) does not.

Task complexity has tremendous relevance for the officer engaged in a potentially hostile encounter. Faced with a range of irrelevant and confusing factors (e.g., darkness, noise, movement, bystanders), the officer must determine whether or not a small and relatively inconspicuous weapon is present. On a reduced scale, our paradigm attempts to simulate this visual and cognitive challenge. The task employs a variety of complex and realistic backgrounds (e.g., parking lots, train stations). By varying backgrounds and suspect poses (e.g., standing, crouching), as well as the timing of stimulus onset, we prevent participants from knowing when or where an object will appear. When the object does appear, it accounts for roughly 0.2% of the visual field. To respond correctly, participants must engage in a careful, controlled search for a small cue amid a complex stimulus array. In contrast to the visually simple tasks typically employed in research on training, training on this relatively complex task may not foster automaticity in the shoot/don't-shoot decision. In our task—as in a police encounter—even highly trained experts may need to fully engage executive control processes to identify the object and execute the appropriate response (Weissman et al., 2002).

If experts are better able than novices to engage control processes, it stands to reason that police officers, whose training and on-the-job experiences routinely force them to identify weapons in complex environments, should make fewer errors in our shoot/don't-shoot task and should show reduced racial bias in those errors (i.e., their expertise should minimize stereotypic errors). This training-based reduction in bias, which we might call a “police as experts” pattern, serves as our primary hypothesis ( $H_1$ ). But control may not entirely eliminate race-based processing. The necessity of a slow, effortful, and controlled search for the object

leaves open the possibility that even experts will inadvertently process racial information. Research suggests that racial cues are often perceived quickly, whether or not the participant intends to do so (Cunningham et al., 2004; Ito & Urland, 2003), and accordingly, a slow visual search for the object should glean racial information. By activating stereotypes, these cues may interfere with the speed of the decision-making process. By virtue of enhanced control, experts may rarely, if ever, shoot an unarmed Black individual; but because even experts must search (slowly) for the object, they are likely to perceive the target's skin color and facial features, triggering relevant stereotypes. Again, experts may effectively override this interference and make an unbiased response (“don't shoot”), but because the weapon judgment is not automatic, the controlled decision to stereotype incongruent targets may still take more time. This leads us to predict a dissociation, such that a target's race may affect the *speed* of the expert's decisions, even though it has no impact on their *accuracy*.

To examine this possibility, the present research extends past work in two critical ways. First, we examine bias in both response times and errors. In past research (e.g., Correll et al., 2002; Payne, 2001), results from these two measures mirrored one another and were characterized as more or less interchangeable. But the measures may capture partially distinct aspects of the decision process. *Latency*—the time necessary for a participant to respond correctly to a given target—should depend on the difficulty of processing the stimulus. The fact that stereotype-incongruent targets (unarmed Black targets and armed White targets) generally produce longer latencies suggests that participants have greater difficulty arriving at a correct decision for these stimuli. Processing difficulty may also influence *error rates*, but errors also reflect the participant's ultimate decision about which response to make. Particularly from an officer's perspective, the distinction between a slow-but-accurate response (e.g., hesitating and then deciding not to fire) and an incorrect response (e.g., shooting an unarmed suspect) assumes great importance.

This research also advances our understanding by comparing police officers with samples of laypeople drawn from the communities those officers serve. Community samples provide a crucial baseline against which we can compare the police. As we have already discussed, one of the most damaging consequences of officer-involved shootings in which a minority suspect is killed is the implication that police inappropriately use race when making the decision to fire. However, given the prevalence of bias in the decision to shoot (which has been documented in all types of people, from White college students to Black community members), how can we interpret the magnitude of any bias we might observe among the police? Inhabitants of the community served by a given police department provide a critical comparison. As members of a common culture, these individuals experience many of the same influences, whether very global (e.g., national broadcast media) or very local (e.g., racial and ethnic composition of the neighborhood, local levels of poverty and crime) in nature. To fully characterize the presence of any bias among police, it is therefore critical to examine bias in the communities they serve. No such comparison is available in existing research. Although we have elaborated the hypothesis that police will demonstrate less bias than the community, particularly with respect to their error rates ( $H_1$ ), we note that the comparison between police and community presents two other possibilities.

Of course, it is also possible that officers will show more pronounced bias than community members ( $H_2$ ) or that police and civilians will show relatively similar patterns of bias ( $H_0$ ). In line with the former hypothesis, Teahan (1975a, 1975b) presented evidence that police departments acculturate White officers into more prejudicial views during their first years on the job. Similarly, the Christopher Commission's investigation into the Los Angeles Police Department's 1991 beating of Rodney King reported that officers who adopted anti-Black attitudes were more likely to be promoted within the department (Christopher, 1998). This ostensible culture of bias may find expression in police officers' relatively high social dominance orientation (Sidanius & Pratto, 1999), reflecting support for the group-based (and race-based) hierarchical structure of society (see Sorenson et al., 1993, for similar conclusions on the basis of police use of force). Given these findings, we might reasonably expect a "police as profilers" pattern, with officers relying heavily on racial information when making their decisions to shoot.

Finally, police officers and community members may show equivalent levels of racial bias in decisions to shoot. Inasmuch as police and community members are subject to the same general cognitive heuristics (Hamilton & Troler, 1986) and sociocultural influences (Devine & Elliot, 1995), the two groups may demonstrate similar patterns of behavior in the video game simulation. This prediction would yield a pattern we might call "police as citizens."

Our primary hypothesis derives from the possibility that practice enables police officers to more effectively exert control over their behavioral choices (relative to untrained civilians). That is,  $H_1$  suggests that officers may more extensively engage in controlled processing operations during the course of the shoot/don't-shoot task. Because of this difference in processing, we predict a divergence between measures of bias that are based on errors and measures that are based on reaction times. By contrast,  $H_2$  and  $H_0$  offer no clear reason to predict differences between officers and civilians in terms of cognitive processing, and (accordingly) they offer no reason to expect a divergence between error-rate and reaction-time measures.

## Study 1

### Method

*Overview.* Three samples of participants completed a 100-trial video game simulation in which armed and unarmed White and Black men appeared in a variety of background images. Participants were instructed that any armed target posed an imminent threat and should be shot as quickly as possible. Unarmed targets posed no threat and should be flagged accordingly by pushing the don't-shoot button, again as quickly as possible. The speed and accuracy with which these decisions were made served as our primary dependent variables, and performance was compared across three samples: officers from the Denver Police Department, civilians drawn from the communities those officers served, and a group of officers from across the country attending a 2-day police training seminar.

*Participants.* For the purposes of law enforcement, the city of Denver is divided into six districts. With the help of the command staff, officers were recruited for this study from four of these districts during roll call. Participation was completely voluntary, and officers were assured that there would be no way to identify individual performance on the task and that the command staff would not be informed of who did and did not participate. Officers were required to complete the simulation during off-duty hours. Our goal was to recruit primarily patrol officers, and, in this effort, we were successful: 84% of the sample listed patrol as their job category. Investigative officers accounted for 9% of the sample, administrative officers for 2% of the sample, with the remaining 5% of the officers from a mixture of other job categories. A total of 124 officers participated in the study (9 female, 114 male, 1 missing gender; 85 White, 16 Black, 19 Latina/o, 3 other, 1 missing ethnicity; mean age = 37.9 years). Each received \$50.

To obtain a companion civilian sample, we enlisted the Department of Motor Vehicles (DMV) office in each of the four districts, recruiting community members to perform the simulation on or around the same days as the police officers. Several of the DMVs were in areas with a high concentration of Spanish-speaking citizens. For these areas, a bilingual research assistant recruited and instructed the participants.<sup>1</sup> A total of 135 civilians participated in the study. Eight participants were dropped from the analyses: 2 because of a computer malfunction and 6 because they had fewer than five correct trials for at least one of the four cells of the simulation design. Thus, the reported results for this sample are based on 127 civilians (51 female, 73 male, 3 missing gender; 39 White, 16 Black, 63 Latina/o, 9 other; mean age = 35.5 years). Each received \$20.

To collect the national police sample, we attended a training seminar for officers. This was one of several seminars that officers voluntarily attend to obtain additional training in some particular area of law enforcement. The seminars are specifically geared for patrol officers, rather than administrative personnel. The sample of officers obtained for this study came from 14 different states, and only 7% worked in some administrative capacity. The remaining job categories included patrol officers (58%), investigative officers (14%), traffic officers (7%), SWAT team members (3%), and a sprinkling of other categories (11%). Although this clearly is not a random national sample of officers, it offers a greater diversity of background than the Denver sample. An announcement regarding the study was made during the seminar, and officers were invited to participate on one of two evenings after the conclusion of the seminar for that day. A total of 113 officers participated in the study (12 female, 100 male, 1 missing gender; 72 White, 10 Black, 15 Latina/o, 13 other, 3 missing ethnicity; mean age = 38.4 years). Each received \$50.

*Video game simulation.* Fifty men (25 Black, 25 White) were photographed in five poses holding one of a variety of objects, including four guns (a large black 9 mm, a small black revolver, a large silver revolver, and a small silver automatic) and four non-guns (a large black wallet, a small black cell phone, a large silver Coke can, and a small silver cell phone). For each individual, we

<sup>1</sup> Many thanks to Alinne Barrera who tirelessly and happily made time in her busy schedule to accompany us on these sojourns at the Denver DMVs.

selected two images, one with a gun and one with an innocuous object, resulting in 100 distinct images (25 of each type: armed White, armed Black, unarmed White, and unarmed Black), which served as the principal stimuli, or targets, in the game. Forty of these images were drawn from previous work (see Correll et al., 2002, for example stimuli). The others were added in an effort to diversify the sample of targets. Using Photoshop, we embedded targets in 20 otherwise unpopulated background scenes, including images of the countryside, city parks, facades of apartment buildings, and so on. Each target was randomly assigned to a particular background, with the restriction that each type of target should be represented with equal frequency in each background.

*Design.* The video game, developed in PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993), followed a  $2 \times 2$  within-subjects design, with Target Race (Black vs. White) and Object Type (gun vs. nongun) as repeated factors (see Correll et al., 2002). On any given trial of the game, a random number (0–3) of preliminary backgrounds appeared in slideshow fashion. These scenes were drawn from the set of 20 original unpopulated background images. Each remained on the screen for a random period of time (500 ms–800 ms). Subsequently, a final background appeared (e.g., an apartment building), again for a random duration. This background was replaced by an image of a target person embedded in that background (e.g., an armed White man standing in front of the building). From the player's perspective, the target simply seemed to appear in the scene. The player was instructed to respond as quickly as possible whenever a target appeared, pressing a button labeled *shoot* if the target was armed and pressing a button labeled *don't shoot* if the target was unarmed. The game awarded points on the basis of performance. Correctly pressing *don't shoot* in response to an unarmed target earned 5 points, but shooting earned a penalty of 20 points; pressing *shoot* in response to an armed target earned 10 points, but pressing *don't shoot* earned a penalty of 40 points (the implication being that the hostile target shot the player). Failure to respond to a target within 850 ms of target onset resulted in a penalty of 10 points. Feedback, both visual and auditory, and point totals were presented at the conclusion of every trial. The game consisted of a 16-trial practice block and a 100-trial test block.

*Procedure.* Officers in the Denver sample were recruited roughly 1 week prior to the study. Volunteers selected a time and date to participate. At the scheduled time, each officer was seated at a small cubicle in a test room equipped with a laptop computer, button box, and headphones. They completed the simulation and questionnaire packet. The measures included in the questionnaire packet are summarized in Table 1. Community members were approached at one of the various DMV locations, and those who agreed to participate completed the simulation using the same equipment as the officers. Community members completed the same questionnaire as the officers (excluding items specific to policing). For the national sample of officers, an announcement was made the first day of the training seminar inviting officers to participate in the study. Officers completed the simulation and questionnaire packet on one of two evenings in a room in the hotel where the conference was held. The equipment was identical to that used for the Denver officers and civilians. Upon completion, all participants were debriefed and thanked.

## Results

*Signal-detection analyses.* We began by examining the accuracy of responses as a function of trial type and sample. Overall, participants responded incorrectly on 4.7% of the trials and timed out on another 4.8% of the trials. Correct and incorrect responses (i.e., excluding timeouts) were used to conduct a signal-detection analysis. Applied to the shooter simulation, signal detection theory (SDT) assumes that armed and unarmed targets vary along some dimension relevant to the decision at hand (e.g., the threat they pose). SDT yields estimates of participants' ability to discriminate between the two types of target (i.e., sensitivity to the presence of a weapon, a statistic called  $d'$ ) and the point on that decision-relevant dimension at which they decide a stimulus is threatening enough to warrant shooting (i.e., the psychological criterion for the decision to shoot, a statistic called  $c$ ). With SDT it is possible to test whether the race of a target affects discriminability and, separately, whether target race affects the decision to shoot. Correll et al. (2002, Study 2) observed no race differences in  $d'$  but found that  $c$  was lower for Black targets than for White targets. That is, participants were equally able to differentiate between armed and unarmed targets regardless of target race, but they used a more lenient threshold—indicating a greater willingness to shoot—when the target was Black rather than White.

We calculated  $d'$ , or the ability to accurately discriminate armed from unarmed targets, once for the White targets and once for the Black targets. We also calculated  $c$ , or the criterion, assessing the threshold for making a *shoot* response separately for Black and White targets.<sup>2</sup> The SDT estimates were submitted to separate  $3$  (Sample: national officers vs. Denver officers vs. Denver community)  $\times 2$  (Target Race: Black vs. White) mixed-model analyses of variance (ANOVAs).

Placement of the criterion for the decision to shoot ( $c$ ) at zero indicates no greater tendency to make a *shoot* response than a *don't-shoot* response. Deviations from zero in a positive direction indicate a bias favoring the *don't-shoot* response, and deviations in a negative direction indicate a bias to shoot. On average (i.e., for both officers and civilians and both Black and White targets), participants demonstrated a bias in favor of the *shoot* response,  $F(1, 361) = 4.68, p = .03$ , but the extent to which this was true depended on sample,  $F(2, 361) = 4.97, p < .008$ . Pairwise comparisons indicated that the community set significantly lower criteria than either officer sample, both  $F_s(1, 361) > 4.12, p_s < .05$ . (All pairwise comparisons were tested with the error term from the full sample.) Indeed, although the mean threshold was significantly below zero for the community sample,  $F(1, 126) = 10.05, p < .002$ , it did not differ from zero for either of the two officer samples, both  $F_s < 1$ , and the two officer samples did not differ from each other,  $F(1, 361) = 1.22, p = .27$ .

It is important to note that the main effect of target race in the placement of the decision criterion was significant,  $F(1, 361) = 5.17, p < .03$ , such that  $c$  was lower when responding to Black

<sup>2</sup>  $c = -0.5 \times (zFA + zH)$ ;  $d' = zH - zFA$ , 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  $n$  = the total number of no-gun and gun trials, respectively) and a maximum of  $1 - (1/2n)$ , to eliminate infinite  $z$ -scores.

Table 1  
*Demographic and Psychological Variables Included in Questionnaire Packet and Their Correlations With Bias in Latencies in Study 1*

Variable	Correlation with bias in latencies		
	National officers	Denver officers	Denver community
Population of city in which officer serves	.31***	—	—
Population of county in which officer serves	.31***	—	—
Violent crime in community served	.20**	-.09	-.05
% African Americans in community served	.21**	.01	.01
% all ethnic minority groups in community served	.22**	-.02	.05
Classroom firearms training	.01	—	—
Firing-range firearms training	.03	—	—
Video firearms training	-.02	—	—
Live firearms training	.02	—	—
Total years on the force	-.09	-.17*	—
Gender (-1 = female; 1 = male)	-.13	-.13	.21**
Ethnicity (-1 = non-White; 1 = White)	-.09	-.14	-.08
Education	.02	.10	-.12
Self-rated liberalism (1)–conservatism (13)	-.04	-.21**	-.06
Thermometer rating (warmth toward White people–warmth toward Black people)	.00	-.02	.03
Thermometer rating (warmth toward White people–warmth toward members of all ethnic minority groups)	.00	.00	-.04
Personal stereotype of Black people as dangerous, violent, and aggressive	.02	.01	.20**
Cultural stereotype of Black people as dangerous, violent, and aggressive	.02	.05	.09
Contact with Black people	.05	-.02	.11
Internal motivation to control prejudice	-.04	.05	-.11
External motivation to control prejudice	.16	-.12	.20**
Discrimination scale	-.13	-.04	.08

*Note.* City and county population have no variance for the Denver police and community samples, and hence no correlation can be computed. Firearms training data were not collected for the Denver officers, nor for the community. *Ns* vary slightly across entries because of missing observations. In the national sample, *ns* vary from 97–113; in the Denver police sample, they vary from 118–123; and in the Denver community sample, they vary from 120–127. Dashes indicate that data were not collected.

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

rather than White targets (see the top half of Figure 1 and the means in Table 2). This discrepancy constitutes bias. Although the omnibus test of the interaction between target race and sample was not significant,  $F(2, 361) = 1.87, p = .16$ , pairwise comparisons indicated a larger target race difference for the Denver community compared with the national officer sample,  $F(1, 361) = 3.67, p = .056$ , other  $F_s < 1.49, ps > .22$ . Racial bias in  $c$  was significant among the Denver community sample,  $F(1, 126) = 5.71, p < .02$ , marginally significant among the Denver officer sample,  $F(1, 123) = 3.28, p = .07$ , and nonsignificant among the national officer sample,  $F < 1$ .

It is informative to examine sample differences in  $c$  separately for the White and Black targets. As is clear from Figure 2, placement of the criterion for the White targets changed very little across the three samples, and in fact neither the omnibus test of sample differences,  $F < 1$ , nor any of the pairwise comparisons, all  $F_s(1, 361) < 1.54, ps > .22$ , revealed a significant difference on this measure. Moreover, the criterion for White targets was not significantly different from zero for any of the three samples, all  $F_s < 1.49, ps > .23$ . That is, neither officers nor community members showed a tendency to favor one response over the other when the target was White. In contrast, the threshold for Black

targets changed substantially and significantly across the three samples,  $F(2, 361) = 7.03, p < .001$ . The criterion was set lowest by the Denver community sample, whose mean  $c$  was both significantly lower than zero,  $F(1, 126) = 15.05, p < .001$ , and significantly lower than either of the two officer samples, both  $F_s(1, 361) > 4.42, ps < .04$ . The Denver officers' mean  $c$  value was also significantly below zero,  $F(1, 123) = 4.04, p < .05$ , and approached a significant difference when compared to the national officer sample,  $F(1, 361) = 2.79, p = .10$ . The national officers' criteria for Black targets did not differ from zero,  $F < 1$ .<sup>3</sup>

<sup>3</sup> In each of the three samples, we tested for moderation of bias in latencies,  $d'$ , and  $c$  by participant ethnicity and gender. Because of the relatively small number of non-White participants, particularly in the officer samples, these analyses compared all non-White participants with White participants. Bias was not moderated by participant ethnicity for any of the samples ( $ps$  ranged from .76 to .11). The only effect of gender was moderation of bias in response times for the community sample. Bias was significantly greater for male than for female community members,  $F(1, 122) = 5.66, p < .02$ , but it is important to note that bias was significant within each sample,  $F(1, 50) = 11.16, p < .002$  for female participants, and  $F(1, 72) = 61.00, p < .001$  for male participants.

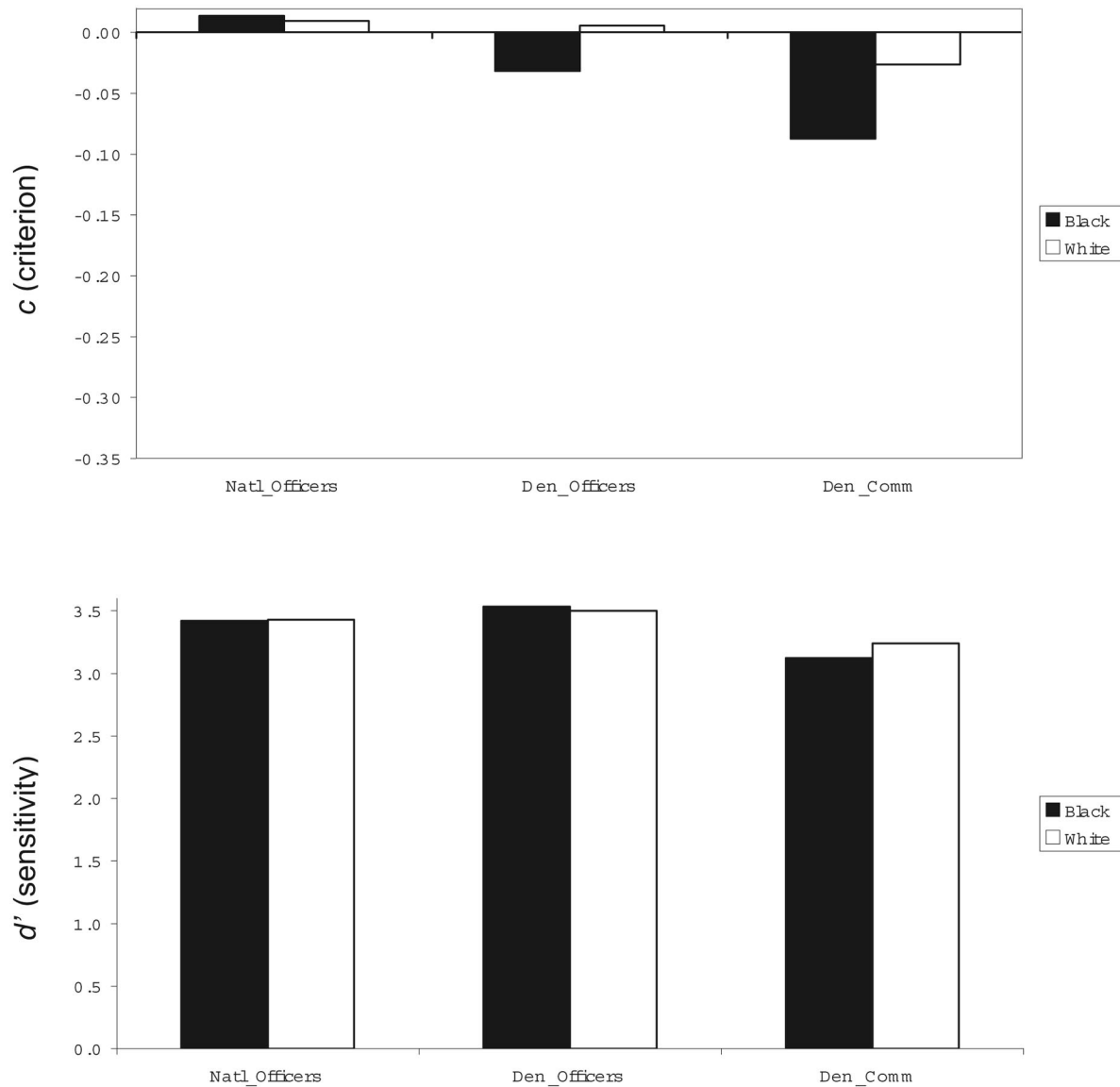


Figure 1. Decision criterion placement ( $c$ ) and sensitivity ( $d'$ ) for Black and White targets as a function of sample (Study 1).

With respect to the analysis of  $d'$ , these data largely replicated previous work, such that target race did not affect participants' ability to discriminate armed from unarmed targets. In other words, the main effect of target race was not significant in the  $d'$  analysis,  $F(1, 361) = 1.12, p = .29$  (see the bottom panel of Figure 1 and Table 2 for all means and standard deviations). However, the main effect of sample was significant,  $F(2, 361) = 11.69, p < .001$ . Pairwise comparisons indicated that both officer samples showed higher discriminability than the community, indicating a greater ability to differentiate armed from unarmed targets, both  $F_s(1, 361) > 11.01, p_s < .001$ . The two officer samples did not differ from one another,  $F(1, 361) = 1.55, p = .21$ . The interaction between sample and race of target was marginally significant,  $F(2, 361) = 2.49, p = .085$ . Pairwise comparisons indicated a significant difference only between the Denver

officers and the Denver community,  $F(1, 361) = 4.63, p < .04$ . The officers showed slightly (but nonsignificant,  $F < 1$ ) greater sensitivity to weapon detection for Black rather than White targets. Among the community,  $d'$  was significantly higher for White targets than for Black targets,  $F(1, 126) = 4.84, p < .03$ .

*Reaction-time analyses.* We next examined reaction times. For each participant, latencies from correct responses were log transformed and averaged separately for each type of target (see Table 2 for means and standard deviations). Averages were analyzed as a function of sample (national officers vs. Denver officers vs. Denver community), target race (Black vs. White), and object type (gun vs. nongun) using a  $3 \times 2 \times 2$  mixed-model ANOVA with repeated measures on the latter two factors. Consistent with past research, we obtained a main effect of object type,  $F(1, 361) = 2,171.27, p < .001$ , such that participants shot armed targets more

Table 2  
Response Time, Sensitivity, and Decision Criterion Means and Standard Deviations for Studies 1 and 2

Variable	Sample											
	National officers				Denver officers				Denver community			
	Black		White		Black		White		Black		White	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Study 1												
Response time												
Gun												
ms	552 <sub>a</sub>		560 <sub>b</sub>		560 <sub>a</sub>		572 <sub>b</sub>		568 <sub>a</sub>		578 <sub>b</sub>	
log transformed mean	6.31	0.07	6.33	0.07	6.33	0.08	6.35	0.07	6.34	0.09	6.36	0.09
No gun												
ms	648 <sub>a</sub>		635 <sub>a</sub>		653 <sub>a</sub>		637 <sub>b</sub>		663 <sub>a</sub>		649 <sub>b</sub>	
log transformed mean	6.47	0.06	6.45	0.06	6.48	0.05	6.46	0.06	6.50	0.07	6.48	0.07
Sensitivity ( <i>d'</i> )	3.42	0.59	3.43	0.50	3.54	0.52	3.50	0.59	3.12	0.78 <sub>a</sub>	3.24	0.76 <sub>b</sub>
Threshold ( <i>c</i> )	.014	0.19	.009	0.21	-.032*	0.18	.006	0.21	-.087*	0.25 <sub>a</sub>	-.026	0.24 <sub>b</sub>
Study 2												
Sensitivity ( <i>d'</i> )					2.39	0.80	2.17	0.73	1.39	0.84	1.47	1.03
Threshold ( <i>c</i> )					-.072	0.30	-.122*	0.31	-.302*	0.33 <sub>a</sub>	-.185*	0.39 <sub>b</sub>

Note. Different row subscripts within each sample indicate a significant Black–White difference at  $p < .05$ . For the decision criterion, means significantly different from zero at  $p < .05$  are indicated with an asterisk.

quickly than they decided to not shoot unarmed targets. The target race main effect was also significant,  $F(1, 361) = 4.90, p < .03$ , such that, overall, responses were very slightly faster to White ( $M = 605$  ms) than to Black targets ( $M = 608$  ms). Moreover, the sample main effect was significant,  $F(2, 361) = 5.36, p < .006$ . Contrasts among the samples indicated that both officer groups responded significantly faster overall than the civilian group,  $F_s(1, 361) > 3.68, ps < .056$ , and the two officer samples did not differ from each other,  $F = 1.86, p = .18$  ( $M_{\text{national officers}} = 597$  ms,  $M_{\text{Denver officers}} = 604$  ms,  $M_{\text{Denver community}} = 613$  ms).

It is important to note that we obtained the Target Race  $\times$  Object Type interaction,  $F(1, 361) = 239.37, p < .001$ . This effect reflects racial bias in decisions to shoot (see Figure 2). Notably, the interaction did not depend on sample,  $F(2, 361) = 1.74, p = .18$ .

Bias was significant for all three samples: for the national sample of officers,  $F(1, 112) = 68.89, p < .001$ , for the Denver officers,  $F(1, 123) = 117.29, p < .001$ , and for the Denver community sample,  $F(1, 126) = 65.29, p < .001$ . Pairwise comparisons among the samples revealed no differences in the magnitude of bias between the community sample and either of the officers samples,  $F_s < 1.17, ps > .28$ , and marginally greater bias among the Denver than national officer sample,  $F(1, 361) = 3.44, p = .065$ .

We further examined the simple effects of target race for each type of object. Again, consistent with previous findings, participants shot armed targets more quickly when they were Black, rather than White,  $F(1, 361) = 74.04, p < .001$ , and they indicated *don't shoot* in response to unarmed targets more quickly when they

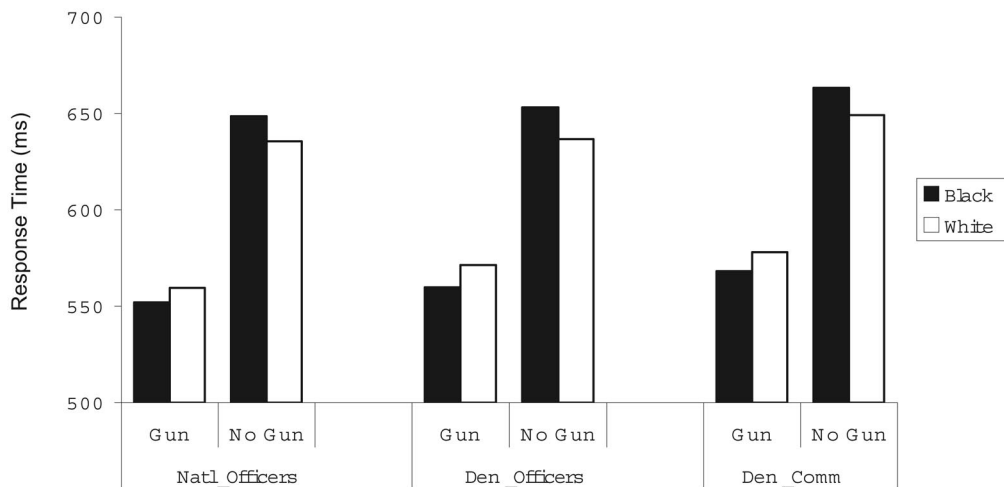


Figure 2. Response times to Black and White armed and unarmed targets as a function of sample (Study 1).



were White, rather than Black,  $F(1, 361) = 177.27, p < .001$ . These simple effects did not depend on sample, both  $F_s < 1, p_s > .39$ , and both of the simple target race effects within object type were significant for each of the three samples, all  $F_s > 15.00$ , all  $p_s < .001$ . Pairwise comparisons for the simple effects among the three samples revealed no significant differences, all  $F_s < 1.85$ , all  $p_s > .17$ .

Summarizing the results thus far, we see that officers and community members differ in the criteria they employ for Black targets. Community members set a lower, more lenient criterion for shooting Black targets than either of the two officer samples. At the same time, officers and community members show similar levels of bias in terms of the speed with which they can correctly respond to targets. We have suggested that, by virtue of their training or expertise, officers may exert control over their behavior, possibly overriding the influence of racial stereotypes. Consistent with the possibility of enhanced control, officers also showed greater sensitivity than did community members to the presence of a weapon, regardless of target race. However, we do not suggest that officers are completely immune to stereotypes. To the extent that a Black target evokes the concept of danger, behavioral control should be difficult. Reactions to targets that violate stereotypic expectancies (i.e., unarmed Black targets and armed White targets) should be slower than reactions to stereotype-congruent targets. If officers' response latencies reflect the magnitude of racial stereotypes, we might expect greater latency bias for officers exposed to stronger environmental associations between Black people and crime. Community characteristics, such as crime rates and the proportion of minority residents, might predict the magnitude of bias among officers in the latencies. It is important to note, however, that if officers can exert control over their behavior, stereotypic associations should not produce greater bias in the SDT criteria they employ. We used the questionnaire data to explore this issue. Because there is very little variance among the Denver officers on these community characteristics (that is, the population of the city and county served by all officers in Denver is the same, and racial makeup across communities varies minimally), the national officer sample affords a more effective test of these possibilities.

**Correlational analyses.** We computed indices of racial bias on the basis of both response times ( $[RT_{\text{unarmed Black target}} - RT_{\text{unarmed White target}}] + [RT_{\text{armed White target}} - RT_{\text{armed Black target}}]$ ), and criteria ( $c_{\text{White}} - c_{\text{Black}}$ ). Higher numbers indicate greater racial bias. We also calculated the effect of target race on discriminability ( $d'_{\text{White}} - d'_{\text{Black}}$ ), with higher numbers representing greater sensitivity for White targets than for Black targets.

We then conducted exploratory analyses of the relationships between each of these indices and the questionnaire measures obtained. We report correlations for all three samples (see Table 1), but again, because the national sample offers greater variability in terms of community demographics, we focus our discussion on that sample. Bias in the response times was positively related to the size (i.e., population) of the city,  $r(97) = .31, p < .003$ , and county,  $r(103) = .31, p < .002$ , in which the officer served (population variables were log transformed to normalize their distributions). This effect suggests that officers in larger communities showed greater bias in the latency measure. In addition, officers' reports of the level of violent crime in their districts predicted bias in response latencies,  $r(111) = .20, p < .03$ , such

that increases in violent crime were associated with greater racial bias. Officers rated violent crime levels with respect to FBI statistics for the average national violent crime rate (500 offenses per 100,000 persons) on a 5-point scale with the endpoints anchored at *much lower than average* and *much higher than average*. Officers were also asked to estimate the ethnic makeup of the communities in which they served. The estimated percentage of both African Americans,  $r(108) = .21, p < .03$ , and ethnic minorities more generally,  $r(108) = .22, p < .03$ , living in the community positively predicted racial bias in the latencies. None of the remaining correlations for the national sample of officers was significant.

Officers serving in districts characterized by a large population, a high rate of violent crime, and a greater concentration of Black people and other minorities showed increased bias in their reaction times. We tentatively suggest that these environments may reinforce cultural stereotypes, linking Black people to the concept of violence. The fact that officers from these urban, violent areas show more pronounced bias in their latencies suggests that stereotypic associations may indeed influence police on some level. But if training enables officers to effectively control their behavior, such stereotypes should not influence their final shoot/don't-shoot decisions. It is interesting that these community demographics, which systematically predicted latency bias, were completely unrelated to bias in the SDT estimates of decision criteria ( $r_s$  ranged from  $-.14$  to  $.13$ , smallest  $p$  value =  $.19$ ). In other words, environmental variables that increased bias in officers' latencies had no effect on the degree of bias in their ultimate decisions.

We also asked participants (community members and officers alike) to complete several measures of stereotyping and prejudice. In the past, we have obtained relationships between bias in response times and an individual's awareness of cultural stereotypes about Black people (Correll et al., 2002, Study 3; Correll, et al., 2007). In the present study, measures of personally endorsed stereotypes did correlate with latency bias for the community members,  $r(123) = .21, p < .05$ , but cultural stereotypes did not. Moreover, in the officers' data, neither of these relationships emerged. It is possible that this difference reflects something special about the relationship between stereotypes and bias among officers, but we suspect that the reason has more to do with the officers' concerns about going "on the record" with regard to their attitudes about race. Despite our assurances of anonymity, several officers were unwilling to complete the measures, and others told us, rather bluntly, that they would not respond honestly to these sensitive questions. We therefore view these items with suspicion, at least for the officer samples.

The effects of target race on the SDT estimates were not related to any of the demographic variables. As null effects, these results are difficult to interpret. They may reflect a true lack of correspondence between demographics and performance, but they may also stem from the relatively low error rates in this task (which likely reduce the reliability of the SDT estimates).<sup>4</sup> Although Black-White differences were unrelated to the questionnaire measures, we did find that the *average* values of both  $d'$  and  $c$  (independent of target race) were correlated with training in simulated building searches. In this type of training, officers interact with actors, some of whom attack the trainee using weapons

<sup>4</sup> We thank an anonymous reviewer for this insight.

equipped with nonlethal ammunition. Police with more extensive training in these encounters were better able to discriminate between armed and unarmed targets, regardless of the race of the target,  $r(113) = .20, p < .04$ , and they tended to set a higher overall criterion in the task,  $r(113) = .17, p < .07$ , reflecting greater reluctance to shoot. It is interesting that no other type of training (e.g., classroom training, firing range, interactive video training) predicted performance in the game. Future researchers should attempt to replicate these correlations, but the results tentatively suggest that live, interactive training provides officers with a chance to hone their skills in a manner that improves performance.

### Discussion

Analyses of the behavioral data showed that the officers' overall performance on the video game simulation exceeded that of the civilians in several ways. First, their response times were faster. On average, officers were simply quicker to make correct shoot/don't-shoot decisions than were civilians. Second, they were better able to differentiate armed targets from unarmed targets. On average (i.e., across White and Black targets),  $d'$  was greater for the officers than for the community sample. Third, whereas the criterion  $c$  for the community was significantly below zero (reflecting a tendency to favor the "shoot" response), officers adopted a more balanced criterion. In fact, not only was the officers' criterion significantly higher than the community's, but the officers' threshold also did not differ significantly from zero. This placement suggests equal use of the "shoot" and "don't shoot" responses.

In terms of bias, the SDT results suggest that officers may show less bias than civilians in their final decisions. Among the community sample, these data revealed a clear tendency to set a lower (i.e., more lenient or "trigger-happy") criterion for Black, rather than White, targets. But this bias was weaker, or even nonexistent, for the officers. The reduction in bias seemed to reflect the fact that, compared with the community members, officers set a higher, more stringent threshold for the decision to shoot Black targets. Placement of the criterion for White targets varied minimally across the three samples.

The response-time data show clear evidence of racial bias for all samples in this study, the 237 police officers and the community members alike. Like college students in previous studies, these individuals seemed to have greater difficulty (indexed by longer latencies) responding to stereotype-incongruent targets (unarmed Black targets and armed White targets), rather than to stereotype-congruent targets. The magnitude of this bias did not differ across the three samples. It is interesting to note that this equivalence emerged in spite of the fact that the civilian sample contained many more ethnic minority members than did the predominantly White police samples. Although any evidence of racial bias among police may be cause for concern, there is certainly nothing in the present data to suggest that officers show greater bias than the people who live in the communities they serve.

We used correlational analyses to examine officers in the national sample, and, of all the variables examined, three predicted bias in reaction times (no variables related to bias in the decision criteria). Each of the relevant variables reflected some aspect of the community the officer served. Bias increased as a function of the community's size, crime rate, and the proportion of Black resi-

dents and other ethnic minority residents. Police in larger, more dangerous and more racially diverse environments are presumably much more likely to encounter Black criminals, reinforcing the stereotypic association between race and crime. By contrast, officers with little exposure to Black people may be less likely to rehearse this association. As a consequence, these officers may experience less stereotypic interference during the video game task.

The results from the signal-detection analysis are particularly provocative. Although police may have difficulty processing stereotype-inconsistent targets (as evidenced by bias in their response times), the SDT results suggest that police do not show bias in their ultimate decisions. That is, the expertise that police bring to a shoot/don't-shoot situation may not eliminate the difficulty of interpreting a stereotype-inconsistent target, but it does seem to minimize the otherwise robust impact of target race on the decision to shoot. Inasmuch as it is the actual decision to shoot (and not the delay in making that decision) that carries life-and-death consequences for the suspect, bias in the criterion may be considered the variable of greatest interest to both the police and the community. However, because of the profound implications of these conclusions, we felt it necessary to replicate these effects. The video game used in Study 1 imposed an 850-ms timeout window. Although this restriction certainly exerts some pressure on participants, it offers them sufficient time to respond correctly on the vast majority of trials. In Study 1, errors and timeouts, together, accounted for only 9.5% of trials. When the total number of errors is so low, idiosyncratic responses to particular targets may dramatically affect the SDT estimates. In Study 2, therefore, we reduced the time window in an effort to increase errors and obtain more stable SDT estimates.

## Study 2

### Method

**Participants.** We returned to one police district in Denver and recruited an additional 33 officers, as well as 52 community members from a nearby DMV, each of whom completed a version of the video game simulation with a more restrictive time window. Several participants experienced great difficulty responding within this limit, producing few errors and a very high number of timeouts. Two officers and 7 civilians had an excessive ratio of timeouts to incorrect trials (more than four timeouts for every error) and were excluded from the analyses. The results do not change substantially if they are included. The final sample included 31 officers (3 female, 26 male, 2 missing gender; 16 White, 6 Black, 4 Latina/o, 3 other, 2 missing ethnicity; mean age = 35.6 years) and 45 community members (20 female, 23 male, 2 missing gender; 14 White, 18 Black, 10 Latina/o, 3 other; mean age = 36.8 years). Officers completed the study while off duty and were paid \$50 in compensation. Community members were paid \$20.

**Video game simulation and procedure.** The video game was identical to that in Study 1, with the exception that the timeout window was set to 630 ms. Participants were instructed to respond as quickly and as accurately as possible, and response latencies longer than 630 ms were penalized with a loss of 20 points. Otherwise, the procedures were identical to those in Study 1.

*Results*

Our goal in reducing the timeout window was to induce a greater number of errors. Our analysis therefore focused on the parameters derived from the signal-detection analysis. Errors were substantially greater in this version of the simulation. Overall, participants made incorrect responses on 16% of the 100 trials and timed out on 17%. We computed sensitivity ( $d'$ ) and the decision criterion ( $c$ ) as in Study 1, using only the correct and incorrect trials (i.e., excluding timeouts). The estimates were analyzed in a Sample (officer vs. civilian)  $\times$  Target Race (Black vs. White)  $2 \times 2$  mixed-model ANOVA, with repeated measures on the latter factor (see Table 2 for means and standard deviations; see also Figure 3).

*Signal-detection analyses.* With respect to the criteria or estimates of  $c$ , we observed that the average criterion was significantly below zero,  $F(1, 74) = 27.06, p < .001$ . In fact, the criteria in Study 2 were lower than those in the first study. Presumably because of the increase in time pressure, participants showed a greater propensity to shoot (compare Figures 1 and 3). More interesting, the location of the criterion depended on sample,  $F(1, 74) = 4.95, p < .03$  (i.e., there was a main effect of sample). Although the mean value of  $c$  was significantly below zero for both the officers,  $F(1, 30) = 4.84, p < .04$  ( $M = -.10$ ), and the community,  $F(1, 44) = 29.38, p < .001$ , ( $M = -.24$ ), it was significantly lower for the latter. Unlike in previous work, the main effect of target race in  $c$  was not

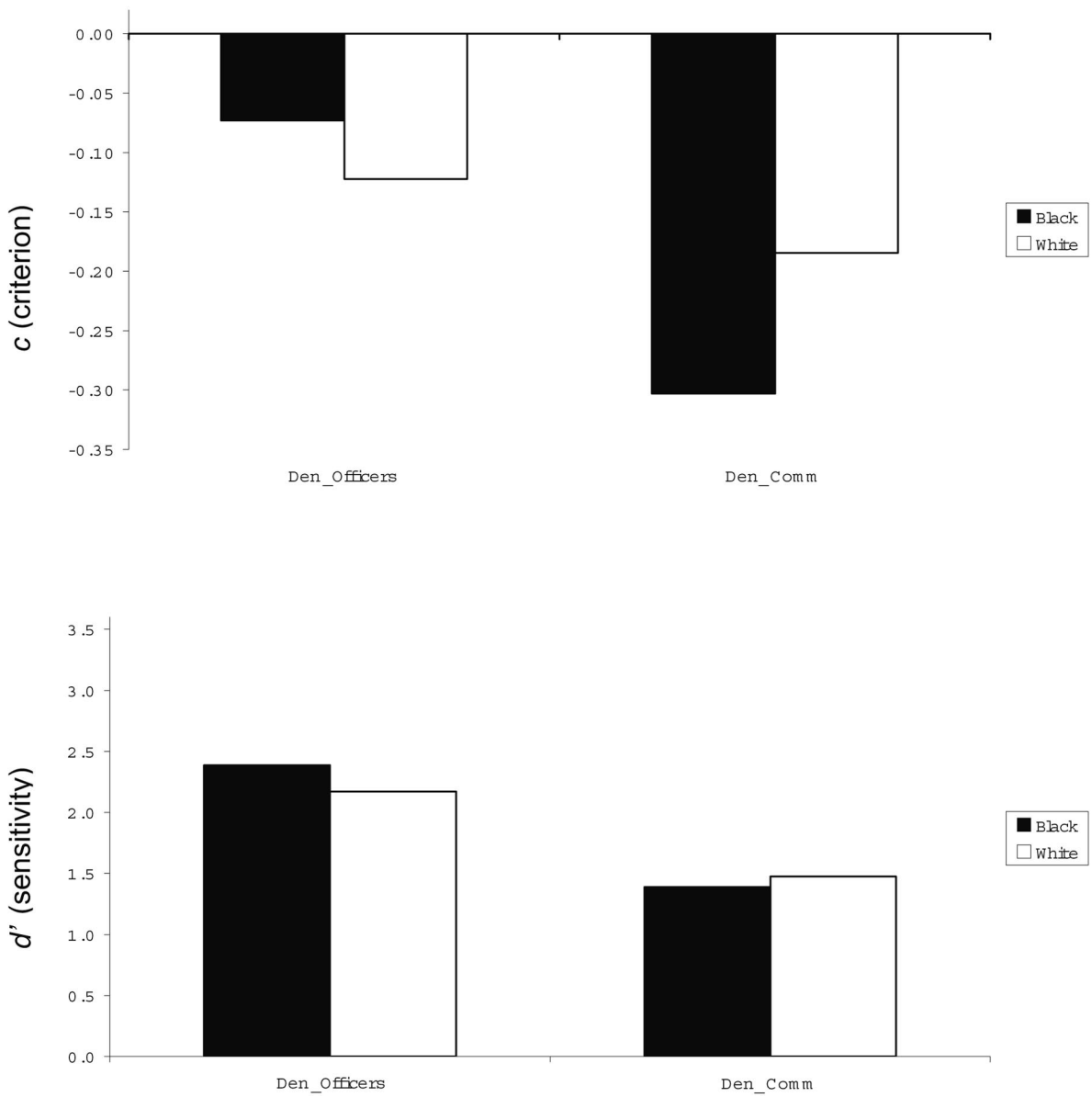


Figure 3. Decision criterion placement ( $c$ ) and sensitivity ( $d'$ ) for Black and White targets as a function of sample (Study 2).

significant,  $F < 1$ , but the Sample  $\times$  Target Race interaction was,  $F(1, 74) = 3.69$ ,  $p = .059$  (see Figure 3). As in Study 1, the community sample set a lower threshold to shoot Black targets than to shoot White targets,  $F(1, 44) = 4.24$ ,  $p < .05$ . Officers, on the other hand, demonstrated no racial bias,  $F < 1$ . Again replicating Study 1, this interaction seems to reflect the fact that the community set a lower threshold for Black targets than did the officers,  $F(1, 74) = 9.74$ ,  $p < .003$ . The two samples did not differ in the placement of their criteria for White targets,  $F < 1$ . It is also interesting to note that all four of the mean  $c$  values in Figure 3 were significantly below zero, all  $t_s < -2.17$ ,  $p_s < .04$ , with the exception of the officers' criterion for Black targets,  $t(30) = -1.36$ ,  $p = .18$ .

Turning to sensitivity, we found that  $d'$  was generally lower in Study 2 than in Study 1, particularly for the community members, suggesting that time pressure impaired discriminability (see Payne, 2001). The main effect of sample was significant,  $F(1, 74) = 21.59$ ,  $p < .001$ . As in Study 1, police officers more effectively discriminated between armed and unarmed targets ( $M = 2.27$ ) than did the community members ( $M = 1.43$ ). The police advantage was evident both for Black targets,  $F(1, 74) = 26.93$ ,  $p < .001$ , and for White targets,  $F(1, 74) = 10.54$ ,  $p < .002$ . There was no overall effect of target race on  $d'$ ,  $F < 1$ , suggesting that participants, in general, were equally able to discriminate White and Black targets. However, target race did interact marginally with sample,  $F(1, 74) = 2.81$ ,  $p < .10$ . Community members were equally sensitive to both White and Black targets,  $F < 1$ , but officers showed marginally greater sensitivity for Black, rather than White, targets,  $F(1, 31) = 3.53$ ,  $p = .07$  (see Figure 3). The results from Study 1 similarly indicated better sensitivity among officers than civilians, particularly for the Black targets.

*Reaction-time analyses.* Previous work has consistently found that reducing the time window eliminates the race-bias effect in response times, presumably because it reduces variance in the latencies (see Correll et al., 2002). Consistent with those findings, bias in response times was not significant on average in Study 2,  $F < 1$ , nor did the magnitude of bias depend on sample,  $F < 1$ .

## Discussion

Like Study 1, Study 2 revealed critical differences between the performance of police officers and that of civilians. These differences emerged both in the participants' ability to discriminate armed from unarmed targets and in the criterion for the decision to shoot. Civilians consistently set a lower threshold for the decision to shoot ( $c$ ) than did the officers, and this difference was particularly evident for Black targets. In both studies, officers showed greater sensitivity ( $d'$ ), and again this tended to be particularly true with Black targets. In sum, then, Study 2 replicated the signal-detection findings of Study 1, and it did so using a paradigm that forced participants to respond very quickly, resulting in a greater number of errors and, so, more stable SDT estimates.

Taken together, the response-time results from Study 1 and the signal-detection results from both Studies 1 and 2 reveal intriguing differences between trained police officers and civilians who live in the communities those officers serve. The latencies suggest that officers and community members both experienced difficulty processing stereotype-incongruent targets. Like community members, police were slower to make correct decisions when faced with an

unarmed Black man or an armed White man. It is important to note, however, that the officers differed dramatically from the civilians in terms of the decisions they ultimately made. Community members showed a clear tendency to favor the *shoot* response for Black targets (relative to both White targets and relative to a neutral or balanced criterion of zero). Police, however, showed no bias in their criteria. Moreover, they showed greater discriminability and a less trigger-happy orientation in general (i.e., for both Black and White targets). These results seem to suggest that expertise improves the outcome of the decision process (increasing sensitivity and reducing the unwarranted tendency to shoot, particularly for Black targets), even though it may not eliminate processing difficulties associated with stereotype-inconsistent targets.

We have suggested that this reduction in bias may reflect the impact of training. In Study 3 we attempted to examine this possibility more systematically by providing practice on the video game task to a sample of undergraduates. On the basis of the results of Studies 1 and 2, we expected that repeated play would improve sensitivity (facilitating discrimination between armed and unarmed targets) and reduce racial bias in the placement of the decision criterion (Plant et al., 2005). But we expected that practice would not reduce bias in response times. Like the officers, participants with more practice on the task should demonstrate improvements in their ultimate decisions in spite of persistent difficulty with the processing of stereotype-incongruent targets.

## Study 3

### Method

*Participants.* Fifty-eight students (29 female, 22 male, 7 missing gender; 40 White, 1 Black, 3 Asian, 3 Latina/o, 1 Native American, 2 Other, 8 missing ethnicity) participated in Study 3 either in partial completion of a course requirement or for \$15 pay. Four additional students were included in the original sample but failed to return for Day 2 and thus are excluded from all analyses.

*Video game simulation and procedure.* In Study 3, participants played the video game twice on each of 2 days separated by 48 hr. At each round of play, they completed an 80-trial shoot/don't-shoot video game, which was essentially the same as the task performed in Study 1. This game again used a timeout window of 850 ms. Thus, the design included four factors: 2 (Day)  $\times$  2 (Round of Play)  $\times$  2 (Race)  $\times$  2 (Object), with repeated measures on all four variables. This design allowed us to examine the effects of repeated play within a day and also to assess whether any improvement in performance would carry over from Day 1 to Day 2.

### Results

We computed SDT estimates and average reaction times for correct responses as in Studies 1 and 2.

*Signal-detection analyses.* We analyzed the SDT estimates as a function of day (1 vs. 2), round of play (1 vs. 2), and target race (Black vs. White) using  $2 \times 2 \times 2$  repeated-measures ANOVAs for both  $c$  and  $d'$ . Analyses of  $c$  revealed that, on average, participants set a lower criterion to shoot for Black targets than to shoot White targets,  $F(1, 57) = 10.76$ ,  $p < .002$ . It is critical, however, that the effect of

race depended on round,  $F(1, 57) = 5.08, p < .03$ , such that bias decreased in the latter round each day. That is, the race difference in the criterion (i.e., bias) was significant at Round 1 on both Day 1,  $t(57) = 2.41, p < .02$ , and on Day 2,  $t(57) = 2.53, p < .02$ . But bias fell to nonsignificant levels at Round 2 on both days: for Day 1,  $t(57) = 0.17, p = .86$ ; for Day 2,  $t(55) = -0.06, p = .95$  (see Figure 4). Moreover, the Round  $\times$  Race interaction did not depend on day,  $F(1, 57) = 0.04, p = .84$ . No other effects in this analysis were statistically significant, all  $F_s(1, 57) < 1.04, p_s > .31$ . As predicted then, practice reduced bias in the decision to shoot, and it did so on each of the two days. It is interesting that there appeared to be no carry over in bias reduction from Day 1 to Day 2. We return to this issue in the Discussion section.

The analysis of sensitivity, or  $d'$ , revealed only a main effect of round,  $F(1, 57) = 7.09, p < .01$ , reflecting greater discriminability during the second game each day. No other effects in this analysis were statistically significant, all  $F_s(1, 57) < 1.06, p_s > .30$  (see Figure 4). As predicted, practice enhanced sensitivity and seemed to have equivalent effects for both Black and White targets. Moreover, the increase in sensitivity occurred each day, and there was no evidence that the increase carried over from Day 1 to Day 2.

*Reaction-time analyses.* Latencies were analyzed as a function of day (1 vs. 2), round of play (1 vs. 2), target race (Black vs. White), and object type (gun vs. nongun) using a  $2 \times 2 \times 2 \times 2$  repeated-measures ANOVA. As usual, we observed a main effect of object,  $F(1, 57) = 409.19, p < .001$ , such that participants

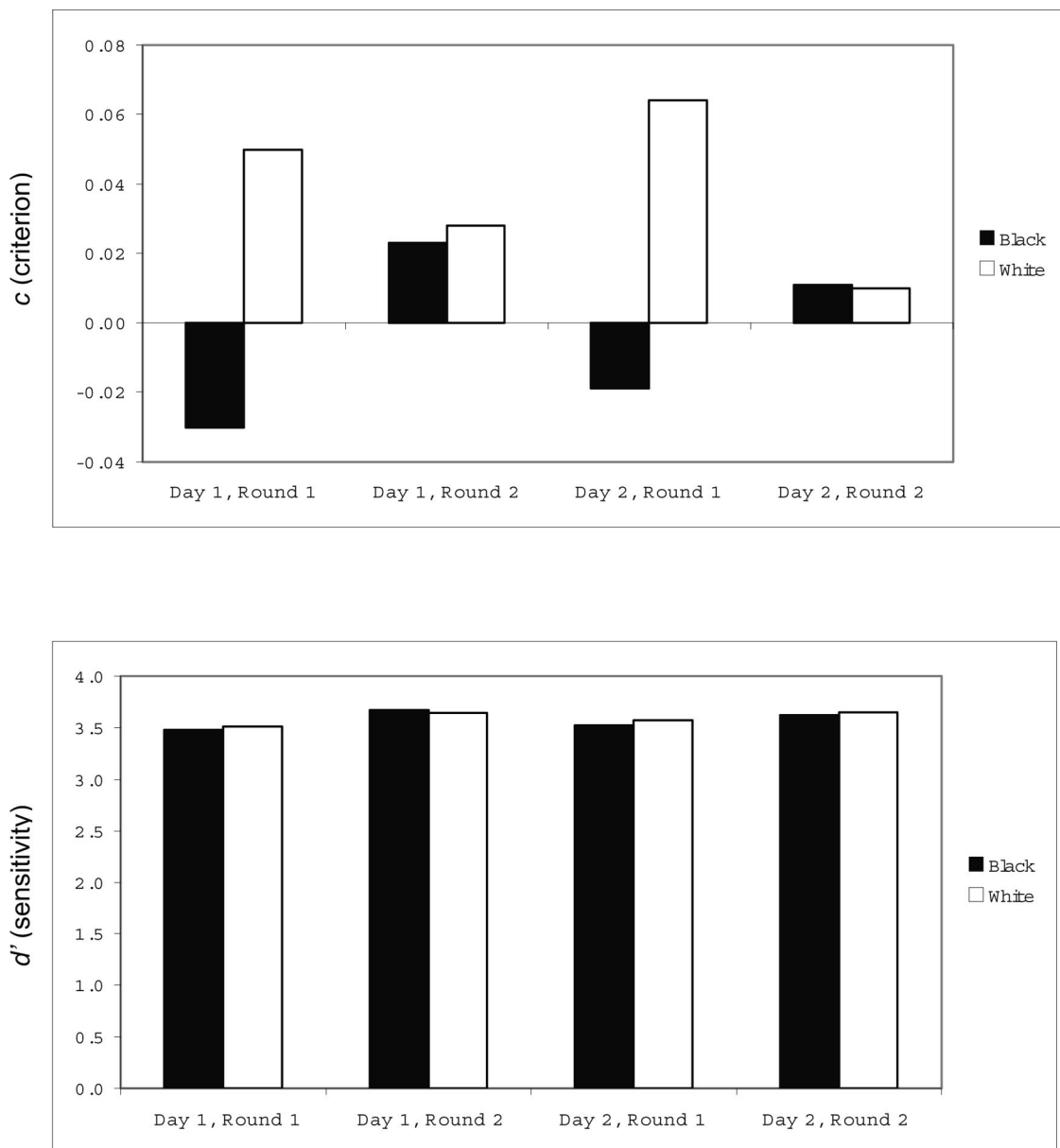


Figure 4. Decision criterion placement ( $c$ ) and sensitivity ( $d'$ ) for Black and White targets as a function of day and round of play (Study 3).

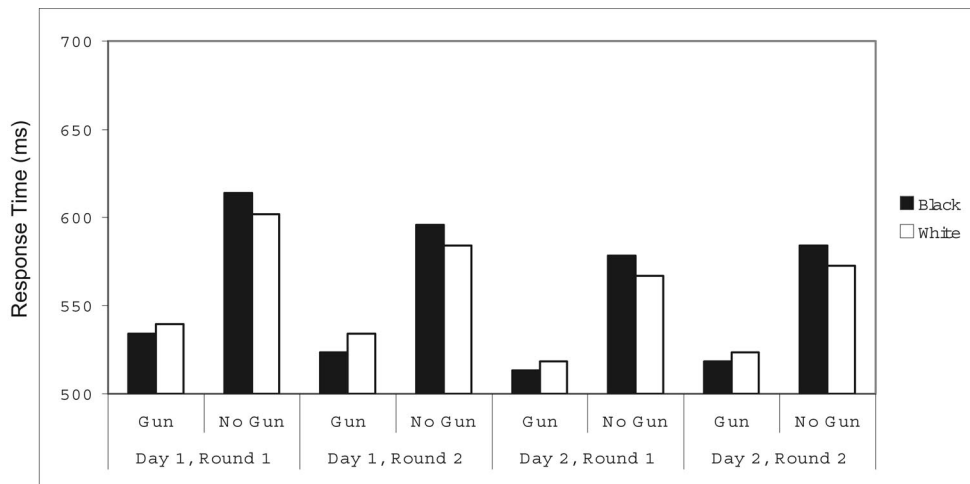


Figure 5. Response times to Black and White armed and unarmed targets as a function of day and round of play (Study 3).

responded more quickly on gun trials than on nongun trials. This effect was qualified by an interaction between target race and object type,  $F(1, 57) = 95.65, p < .001$ , representing significant racial bias. Our primary concern, however, involved the degree to which this pattern changed as participants gained experience with the task. Most interesting, from our perspective, was the question of whether repeated play altered the magnitude of racial bias in the speed with which participants could make shoot/don't-shoot decisions. In stark contrast to the SDT results, bias in reaction times did not change as a function of round: The three-way interaction was not significant,  $F(1, 57) = 0.01, p = .93$ . Similarly, neither the Day  $\times$  Race  $\times$  Object three-way interaction,  $F(1, 57) = 0.01, p = .92$ , nor the Round  $\times$  Day  $\times$  Race  $\times$  Object four-way interaction was significant,  $F(1, 57) = 0.00, p = .95$ . In essence, the magnitude of this bias did not change over the course of the study. Further, latency bias was significant in both Round 1,  $F(1, 57) = 33.76, p < .001$ , and Round 2,  $F(1, 57) = 28.52, p < .001$ , on Day 1, as well as Round 1,  $F(1, 57) = 27.04, p < .001$ , and Round 2,  $F(1, 57) = 17.14, p < .001$ , on Day 2 (see Figure 5).<sup>5</sup> So although practice decreased racial bias in the decision criteria and improved overall discriminability (as shown by the SDT analyses), practice did not attenuate racial bias in reaction times.

### Discussion

Participants in Study 3 showed a number of changes as a function of practice. Most important, practice with the task reduced SDT bias and increased sensitivity to the presence or absence of a weapon. Practice did not, however, affect the magnitude of racial bias in latencies. Across repeated plays of the video game simulation, these developing "experts" continued to struggle with the stereotype-incongruent targets, responding more slowly on incongruent (compared with congruent) trials.

The effects of training observed in this study with a sample of undergraduates largely replicate the differences observed between police officers and civilians in Studies 1 and 2. Undergraduates in the initial round of Study 3, like members of the Denver community, showed bias both in latencies and in their criteria for the

decision to shoot. These effects were evident on both Day 1 and Day 2. After receiving practice on the shoot/don't-shoot simulation task, however, bias in the placement of the criterion diminished, but bias in reaction times did not change. As a consequence of this shift, our "expert" participants began to look less like community members and more like police officers.

However, a single round of practice with our video game task (which takes roughly 12 min–15 min) differs dramatically from the training that police receive. As noted above, Denver police recruits spend approximately 72 hr in weapons training during their time at the academy. This extended, in-depth practice likely results in much greater consolidation of the skills necessary to exert control over their behavior than did the minimal practice afforded to participants in Study 3. Consistent with this, participants in Study 3 showed pronounced within-day improvements (reductions in bias and increases in discriminability), but they showed no evidence that this training carried over from Day 1 to Day 2. Upon entering the lab on Day 2 (48 hr after the Day 1 session), partic-

<sup>5</sup> A number of less theoretically interesting effects that did not involve race and object were present in this analysis. Overall, participants were faster on Day 2 than Day 1,  $F(1, 57) = 46.94, p < .001$ , marginally faster at Round 2 than Round 1,  $F(1, 57) = 3.40, p = .07$ , and the Day  $\times$  Round interaction was significant,  $F(1, 57) = 11.76, p < .002$ , such that the Round 1 to Round 2 decrease in mean latencies was really only present on Day 1. (It is interesting that this increase in speed again mirrors sample differences between the police and community participants in Studies 1 and 2.) The object main effect (faster times to gun trials) was qualified by a number of interactions. The difference in gun versus no-gun trials was greater on Day 1 than Day 2,  $F(1, 57) = 15.69, p < .001$ , for the Day  $\times$  Object interaction, greater at Round 1 than Round 2,  $F(1, 57) = 6.64, p < .02$ , for the Round  $\times$  Object interaction, and the shift from Round 1 to Round 2 was really only present on Day 1,  $F(1, 57) = 4.16, p < .05$ , for the Day  $\times$  Round  $\times$  Object interaction. All of these effects reflect accelerations in classification speed (for all responses or for the particularly slow no-gun responses). This acceleration is most pronounced at early stages of the study and weakens over time, presumably because of a floor effect.

ipants behaved like novices. On Round 1 of their second day, they demonstrated racial bias in both response times and SDT criteria. With additional training on Day 2, this bias dropped once again. But the reemergence of bias in Round 1 of Day 2 suggests more extensive training is necessary if participants are to more permanently overcome bias in behavioral responses. The fact that police officers in Studies 1 and 2 showed no SDT bias during their initial performance on the video game task may be a testament to their training and expertise.

### General Discussion

We began this research with two primary goals: examining police officers in a first-person shoot/don't-shoot task and comparing their performance with that of a community sample. This investigation assessed overall proficiency and the role that a target's race plays in the decision-making process. Police differed from the community members in terms of several critical variables. On average (ignoring target race), the officers clearly outperformed the community sample. They were faster to make correct responses; they were better able to detect the presence of a weapon (as measured by  $d'$ ); and they set a significantly higher criterion ( $c$ ) for the decision to shoot, indicating a less "trigger-happy" orientation.

Most important for our hypothesis, the officers also differed from the community sample in the role that a target's race played in the placement of SDT criteria for the decision to shoot. This difference primarily affected Black targets. When the target was White, all of the samples (Denver community, Denver police, and national police) set a relatively high criterion, and none of the samples differed from one another. But when the target was Black, the community set a significantly lower (more trigger-happy) criterion than the officers. This was true both in Study 1, which used a relatively long timeout window, and in Study 2, in which the timeout window was substantially reduced (yielding much higher error rates).

In spite of the fact that police showed minimal bias in the SDT analysis, the officers were similar to the community sample (and to literally hundreds of past participants in this paradigm) in the manifestation of robust racial bias in the speed with which they made shoot/don't-shoot decisions. Accurate responses to targets congruent with culturally prevalent stereotypes (i.e., armed Black targets and unarmed White targets) required less time than did responses to stereotype-incongruent targets (i.e., unarmed Black targets and armed White targets). Evidence of bias in response latencies was consistent and robust across all of the samples examined in Study 1: the national sample and the Denver sample of police officers, as well as the Denver community sample, drawn from the neighborhoods that the Denver officers serve.

The results from Study 3, in which we trained novice college students on the task, revealed similar effects. Across two rounds of play, student participants showed a significant decrease in racial bias, as measured by the decision criterion, accompanied by an increase in sensitivity. But they showed no change in the magnitude of bias as measured by response latencies. An identical pattern was obtained when students returned for a second day, during which they again completed two rounds of the video game task. In the first round of play, student performance mirrored that

of the community; By Round 2, it mirrored that of the police officers.

The performance of the officers and the expert students in these studies raises an important set of questions about the processes that differentiate bias in response times from bias in the threshold to shoot. Typically, errors and latencies follow a similar pattern, such that greater difficulty on a given trial increases both response time and the likelihood of a mistake, as observed in the performance of community members and novice college students. The officers and experts, by contrast, showed clear bias in latencies, but target race had no impact on their ultimate decisions.

To the extent that longer latencies reflect difficulty, the persistent bias in reaction times suggests that even experts have some trouble processing stereotype-incongruent targets. The visual complexity of the stimuli may essentially require participants to engage in an effortful, serial search for relevant information about the object (Shiffrin & Schneider, 1977). At the same time, the salience and automaticity that generally characterize psychological processing of racial cues (Cunningham et al., 2004; Ito & Urland, 2003) suggest that—during the course of that search—participants are likely to encode target race. In combination with tenacious racial stereotypes (e.g., Devine & Elliot, 1995), race-based processing may impede responses to counterstereotypic targets. In line with this possibility, Study 1 showed that officers from urban, high-crime, predominantly minority districts (environments likely to reinforce stereotypes about Black people) showed greater racial bias in their latencies.

For officers (and, temporarily, for trained undergraduates), however, the stereotypic interference ended with reaction times. The bias evident in their latencies did not translate to the decisions they ultimately made. This separation of effects may reflect the officers' ability to override automatic associations (Kunda & Spencer, 2003), perhaps as a function of their training and expertise. Police (with extensive training) and "expert" undergraduates (with minimal training) were able to reduce bias in the SDT criteria for Black and White targets. Were these individuals able to avoid snap judgments on ambiguous trials, such as those posed by counterstereotypic targets, and wait for a more complete understanding? Such a delay when responding to difficult-to-process counterstereotypic targets would presumably yield bias in reaction times (consistent with the data). At the same time, it would minimize bias in the decision criteria and increase overall accuracy. Anecdotally, this explanation matches officers' intuitions about the process. In a conversation about the effects reported here, one officer stated that the findings "make sense" because police are trained to hold their fire if they are uncertain – to wait for greater clarity.

The possibility that expertise and practice enhance control resonates with research beyond the realm of racial stereotyping. Green and Bavelier (2003) have shown that practice with visually complex video games enhances visual attention (but practice with visually simple games does not). And, although practice on a simple decision task generally promotes automaticity (Bush et al., 1998; Shiffrin & Schneider, 1977), practice on more complicated interference tasks or on challenging working-memory tasks can actually increase control (Olesen et al., 2004; Weissman et al., 2002). On the basis of functional magnetic resonance imaging, these studies show that extended practice on difficult tasks leads to increased activation of the medial and middle frontal gyri—areas

associated with control-based conflict resolution and top-down, rule-based processing. We suggest, then, that police training and on-the-job experience in complex encounters may allow officers to more effectively exert executive control in the shoot/don't-shoot task, essentially overriding response tendencies that stem from racial stereotypes. As noted above, the correlational analyses from Study 1 identified several environmental factors that were associated with increases in *latency* bias for the officers (i.e., serving in urban, high-crime, and predominantly minority districts). It is interesting to note that these same variables had no impact on the SDT criteria the officers used.

We do not want to suggest that the minimal training provided in Study 3 parallels the sort of training that police officers receive. However, the possibility that police function as highly trained subjects is intriguing. In the current research, evidence for this possibility relies on cross-sectional comparisons (Studies 1 and 2) and on parallels between samples that differ in numerous ways (i.e., the "expert" students in Study 3 and the police officers). It would be informative to follow police recruits as they enter the academy, as they receive training, and as they cope with their first years of patrol duty. We have begun data collection on such a project. At present, we have data from 39 recruits in the first weeks of training at the police academy (prior to any weapons training). It is striking that these recruits show statistically significant racial bias in both reaction times and in the decision criteria. Upon entering the academy, then, recruits behave very much like the community samples (Studies 1 and 2) and the novice student sample (Study 3): They set a lower criterion for Black targets than for White targets. These data are entirely consistent with the possibility that the reduction in SDT bias among police officers represents an expertise effect. These data also argue against the suggestion that police academies or departments indoctrinate their members into a culture of anti-Black sentiment (Teahan, 1975a), at least with respect to the sort of judgments studied here.

We must note that our results are only partially consistent with prior work. Consistent with Eberhardt et al. (2004), we found that officers orient more quickly to Black people when processing danger-related stimuli. With respect to reaction times, our results (like theirs) suggest a bias in attentional focus and processing. But our data are not consistent with those of Plant and Peruche (2005), who found that officers showed racial bias in the SDT criteria for the decision to shoot. Although these officers learned to eliminate bias over the course of the study, the presence of the initial bias is inconsistent with our results. Officers in the current studies never showed significant evidence of bias.<sup>6</sup>

This partial correspondence may stem from a variety of factors. We explore two. First, Plant and Peruche (2005) sampled 50 officers from Florida; in Study 1, we sampled 237 officers from Colorado and 14 other states. It is possible that the differences between our findings reflect regional differences between Florida and other areas of the country. Second, it is possible that the results reflect differences between the paradigms employed. Plant and Peruche's stimuli are, arguably, further removed from the training and experience of police officers than are the stimuli presented in our simulation. Plant and Peruche presented Black and White male faces on which objects (e.g., a gun or wallet) had been superimposed. Our stimuli involve full-body images of men holding guns and other objects. These images are embedded in scenes, such as parks or cityscapes. To the extent that our stimuli more closely

mirror police training (e.g., Firearms Training System or firing range encounters) and on-the-job experiences, an officer's expertise should be more likely to generalize to our task. To the extent that Plant and Peruche's paradigm is less similar to the officers' previous experiences, their participants may have had to learn what was, in essence, a novel task.

As we discussed in the introduction, sociologists have studied the question of racial bias in police shootings for many years. The sociological literature provides a rich, if complicated, context in which to view the results of the current studies. One account that has received substantial attention is that police shoot Black suspects more often than White suspects, per capita, because Black people are disproportionately likely to be involved in crime (particularly violent crime). The Department of Justice (2001) report shows that, just as Black suspects are five times more likely than White suspects to die at the hands of police, police officers are five times more likely to die at the hands of a Black suspect than a White suspect. In a similar vein, Reisig et al. (2004) found that the use of nonlethal force (which seems to depend on suspect race) may actually reflect race-based differences in the suspect's propensity to resist arrest or engage in belligerent behavior toward officers. It is the suspect's hostility, they argue—not race—that prompts a hostile response from the officer. And Inn et al. (1977) report that the number of Black suspects shot by police is proportionate to the number of Black suspects arrested. They tentatively conclude that it is the prevalence of criminal activity among Black people that drives the differential shooting rates. (The authors note, however, that arrest rates themselves may reflect biases held by the police and thus do not provide a perfect standard of comparison.) In line with this reasoning, in Study 1, officers from the national sample who reported working in communities with (a) high levels of violent crime and (b) high proportions of minority residents showed particularly strong patterns of bias in their latencies. Did their experiences with minority suspects foster associations that made counterstereotypic trials particularly difficult to process?

The situation is almost certainly more complex. It is clear from the analysis of Study 1 that officers serving in heavily (more densely) populated communities also showed greater anti-Black bias in their reaction times. In combination, these variables seem to suggest that racial bias in the decision to shoot may reflect the disproportionate representation of Black people (and perhaps other ethnic minority groups) in low-income, poverty-stricken, and high-crime areas. Geller (1982) and Smith (2004) presented evidence that a greater number of police shootings occur in disadvantaged neighborhoods and that members of ethnic minorities are more likely to be killed in these incidents. Using regression models to predict officer-involved shootings, Terrill and Reisig (2003) showed that, once neighborhood risk is taken into account, the

---

<sup>6</sup> In light of Plant and Peruche's (2005) findings, we explored the possibility that police officers in the current studies showed a decrease in bias over the course of the shooter task. To examine this possibility, we reanalyzed the data from Studies 1 and 2, separating the 100 trials into two 50-trial blocks and analyzing SDT estimates (both *c* and *d'*) as a function of sample, target race, and block (first half vs. second half). Neither three-way interaction was significant, and controlling for block did not alter the findings reported in the text. These data provide no evidence that police showed less bias than community members because they were better able to improve their performance over the course of the task.



effect of suspect race or ethnicity is no longer statistically reliable. This research builds on the *ecological contamination hypothesis*, first advanced by Werthman and Piliavin (1967), which suggests that the reputation of a neighborhood distorts perceptions of its inhabitants. To the extent that a community is seen as a “bad area,” police may perceive the individuals who live there (or anyone they happen to encounter there) as a potential threat. If members of minorities are more likely to live and spend time in disadvantaged neighborhoods (Sidanius & Pratto, 1999), they may also be more likely to fall victim to this context-based contamination. As a consequence, police may be more likely to shoot a Black suspect because of the context in which the encounter occurs, not because of racial bias, per se (Fyfe, 1981). In an interesting wrinkle of this argument, Sampson and Raudenbush (2004) conducted an extensive investigation of the factors that predict perceived community disorder—the causal variable proposed by ecological contamination. They found that the mere presence of Black people in a community is sufficient to evoke the perception of disadvantage. That is, controlling for objective factors (e.g., prevalence of graffiti, broken windows, and abandoned buildings), the greater the number of Black people living in an area, the greater the disorder perceived by both Black and non-Black citizens. If Black people evoke the perception of neighborhood disadvantage, they may experience harsher treatment by police—not because the police are biased to treat Black people in a hostile fashion, but because Black neighborhoods seem more threatening.

The data presented here suggest that, although police officers may be affected by culturally shared racial stereotypes (i.e., showing bias in their response times), they are no more liable to this bias than are the people who live and work in their communities. Further, at least on the simulation used here, the officers’ ultimate decisions about whether or not to shoot are less susceptible to racial bias than are the decisions of community members. The data suggest that the officers’ training and/or expertise may improve their overall performance (yielding faster responses, greater sensitivity and reduced tendencies to shoot) and decrease racial bias in decision outcomes. We feel that this research represents a valuable melding of basic social psychological processes with an issue of great importance to our society. By examining the influence of race in the automatic processing of danger-related stimuli, and the capacity of expertise to moderate this effect, these findings touch on a topic of great interest to social psychologists, sociologists, police, and community groups, alike. The investigation of racial bias in police use of force presents a unique opportunity to apply experimental social psychological methods to an issue that is vital to the members of increasingly diverse neighborhoods and communities.

## References

- Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of race bias. *Psychological Science, 15*, 88–93.
- Birnboim, S. (2003). The automatic and controlled information-processing dissociation: Is it still relevant? *Neuropsychology Review, 13*, 19–31.
- Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging—Validation study with functional MRI. *Human Brain Mapping, 6*, 270–282.
- Christopher, W. (1998). *Report of the Independent Commission on the Los Angeles Police Department*. Darby, PA: DIANE Publishing.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271.
- 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, 83*, 1314–1329.
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007). *Stereotypes and racial bias in the decision to shoot*. Manuscript submitted for publication.
- Correll, J., Urland, G. L., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology, 42*, 120–128.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science, 15*, 806–813.
- Department of Justice. (2001). *Policing and homicide, 1976–98: Justifiable homicide by police, police officers murdered by felons* (NCJ 180987). Washington, DC: Bureau of Justice Statistics.
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 21*, 1139–1150.
- Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology, 87*, 876–893.
- Fyfe, J. J. (1981). Who shoots? A look at officer race and police shooting. *Journal of Police Science and Administration, 9*, 367–382.
- Geller, W. A. (1982). Deadly force: What we know. *Journal of Police Science and Administration, 10*, 151–177.
- Green, C. S., & Bavelier, D. (2003, May 29). Action video game modifies visual selective attention. *Nature, 423*, 534–537.
- Greenwald, A. G., Oakes, M. A., & Hoffman, H. G. (2003). Targets of discrimination: Effects of race on responses to weapons holders. *Journal of Experimental and Social Psychology, 39*, 399–405.
- Hamilton, D. L., & Trolie, T. K. (1986). Stereotypes and stereotyping: An overview of the cognitive approach. In J. F. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 127–163). San Diego, CA: Academic Press.
- Inn, A., Wheeler, A. C., & Sparling, C. L. (1977). The effects of suspect race and situation hazard on police officer shooting behavior. *Journal of Applied Social Psychology, 7*, 27–37.
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616–626.
- Jansma, J. M., Ramsey, N. F., Slagter, H. A., & Kahn, R. S. (2001). Functional anatomical correlates of controlled and automatic processing. *Journal of Cognitive Neuroscience, 13*, 730–743.
- Kelly, A. M. C., & Garavan, H. (2004). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex, 15*, 1089–1102.
- Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theory of stereotype activation and application. *Psychological Bulletin, 129*, 522–544.
- MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition, 26*, 201–211.
- MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 126–135.

- Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*, 75–79.
- 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*, 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 and Social Psychology*, *38*, 384–396.
- Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, *16*, 180–183.
- Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental and Social Psychology*, *41*, 141–156.
- Reisig, M. D., McCluskey, J. D., Mastrofski, S. D., & Terrill, W. (2004). Suspect disrespect toward the police. *Justice Quarterly*, *21*, 241–268.
- Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of "broken windows." *Social Psychology Quarterly*, *67*, 319–342.
- Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Smith, B. W. (2004). Structural and organizational predictors of homicide by police. *Policing: An International Journal of Police Strategies and Management*, *27*, 539–557.
- Sorenson, J. R., Marquart, J. W., & Brock, D. E. (1993). Factors related to killings of felons by police officers: A test of the community violence and conflict hypotheses. *Justice Quarterly*, *10*, 417–440.
- Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, *18*, 643–662.
- Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review*, *37*, 513–548.
- Teahan, J. E. (1975a). A longitudinal study of attitude shifts among black and white police officers. *Journal of Social Issues*, *31*, 47–56.
- Teahan, J. E. (1975b). Role playing and group experience to facilitate attitude and value changes among Black and White police officers. *Journal of Social Issues*, *31*, 35–45.
- Terrill, W., & Reisig, M. D. (2003). Neighborhood context and police use of force. *Journal of Research in Crime & Delinquency*, *40*, 291–321.
- Weissman, D. H., Woldorff, M. G., Hazlett, C. J., & Mangun, G. R. (2002). Effects of practice on executive control investigated with fMRI. *Cognitive Brain Research*, *15*, 47–60.
- Weitzer, R., & Tuch, S. A. (2004). Race and perceptions of police misconduct. *Social Problems*, *51*, 305–325.
- Werthman, C., & Piliavin, I. (1967). Gang members and the police. In D. Bordua (Ed.), *The police: Six sociological essays*. New York: Wiley.

Received May 1, 2006

Revision received September 17, 2006

Accepted September 28, 2006 ■