J Exp Criminol (2013) 9:189–212
DOI 10.1007/s11292-012-9163-y

Results from experimental trials testing participant responses to White, Hispanic and Black suspects in high-fidelity deadly force judgment and decision-making simulations

Lois James · Bryan Vila · Kenn Daratha

Published online: 23 October 2012

Abstract
Objective Advance the methodological techniques used to examine the influence of suspect race and ethnicity on participant decisions to shoot in an experimental setting.
Methods After developing and testing a novel set of 60 realistic, high definition video deadly force scenarios based on 30 years of official data on officer-involved shootings in the United States, three separate experiments were conducted testing police (n=36), civilian (n=72) and military (n=6) responses (n=1,812) to the scenarios in high-fidelity computerized training simulators. Participants' responses to White, Black and Hispanic suspects in potentially deadly situations were analyzed using a multi-level mixed methods strategy. Key response variables were reaction time to shoot and shooting errors.
Results In all three experiments using a more externally valid research method than previous studies, we found that participants took *longer to shoot Black suspects* than White or Hispanic suspects. In addition, where errors were made, participants across experiments were more likely to shoot unarmed White suspects than unarmed Black or Hispanic suspects, and were more likely to *fail to shoot* armed Black suspects than armed White or Hispanic suspects. In sum, this research found that participants displayed significant bias *favoring* Black suspects in their decisions to shoot.
Conclusions The results of these three experiments challenge the results of less robust experimental designs and shed additional light on the broad issue of the role that status characteristics, such as race and ethnicity, play in the criminal justice

Research supported by the Defence Advanced Research Projects Agency (109279–001), the National Institute of Justice (2008-IJ-CX-0015), and the Office of Naval Research (DURIP#N000149810802)

L. James (✉) · B. Vila · K. Daratha
Washington State University, 412 E. Spokane Falls Blvd., Suite 702, PO Box 1495, Spokane, WA 99210-7810, USA
e-mail: lois_james@wsu.edu

B. Vila
e-mail: vila@wsu.edu

K. Daratha
e-mail: kdaratha@wsu.edu

system. Future research should explore the generalizability of these findings, determine whether bias favoring Black suspects is a consequence of administrative measures (e.g., education, training, policies, and laws), and identify the cognitive processes that underlie this phenomenon.

Keywords Police · Decision making · Suspect race · Suspect ethnicity

Introduction

Concern that the status characteristics of individuals affect how they are treated within the criminal justice system has generated substantial theorizing and research in the social and behavioral sciences during the past several decades. The status characteristic that has received the most attention is race. Scholars have asserted that race plays a major role in how individuals are treated at every phase of the justice system, including their treatment by the police. One specific policing issue that has received a great deal of attention is the use of deadly force. Research indicates substantial disparities in the application of deadly force by the police (Geller 1982; Brown and Langan 2001). Study after study has shown that Blacks are shot by the police at rates higher than expected from their representation in the population (Robin 1963; Takagi 1974; Kobler 1975; Sherman 1979; Meyer 1980; Sorenson et al. 1993; Jacobs and O'Brien 1998).

Police are expected to use deadly force as a last resort, and even then, only when it is within the confines of the law. Police in the United States are justified to use deadly force under two conditions; (1) to prevent loss of life, or (2) to stop the flight of violent fleeing felons (Tennessee v Garner, 1985). Officer-involved shootings risk lives, and the consequences of these encounters can devastate communities. The consequences of police shootings tend to be particularly grave when the subject of police fire is a minority group member. High profile police shootings of unarmed Black men such as Amadou Diallo (Cooper 1999) fuel the belief that police officers are influenced by suspect race when they make split-second decisions about whether a suspect is armed and poses a legitimate risk to others. Riots over fatal shootings or misuse of deadly force against Black males have occurred in alarming numbers over the last 100 years, for example Cincinnati, 2001 (Deakin 1988). Public concern is not limited to the treatment of Blacks, but also extends to the treatment of ethnic minorities. For example, the shooting of Latino journalist Ruben Salazar by police with a tear gas missile fuelled riots in East Los Angeles in 1970 (Lopez 2011). Recently, it also has been suggested that police officers are racially biased against each other. In 2010, the New York State Task Force on Police-on-Police Shootings reported that, during the past 15 years, 10 of the 14 police officers who have been killed in mistaken-identity, police-on-police shootings have been "people of color" (Stone et al. 2010, ii).

Research suggests that Blacks are the subjects of police deadly force disproportionate to their representation in the population (Geller 1982; Brown and Langan 2001). Although no comprehensive nationwide data on police shootings in the United States exist, the FBI supplemental homicide reports suggest that, during the 23-year period from 1976 to 1998, Blacks were approximately four times more likely to be

shot by officers than Whites (U.S. Department of Justice 2001). The statistics on police use of deadly force against ethnic minority suspects such as Hispanics are less clear, in part due to criminal justice agencies not always using the same ethnic categories in reports (Walker 2004). For example, Hispanics may be characterized as "Hispanic White" or "Hispanic Black." For the purposes of this research "Hispanic" refers to "Hispanic White." Research suggests that Hispanics, who make up approximately 16 % of the population, are shot by the police about as frequently as Whites, who make up approximately 70 % of the population (U.S. House Committee on the Judiciary, Subcommittee on Criminal Justice 1989; James and Pasquale-Styles 2009). This suggests that Hispanics, like Blacks, are more likely to be shot by the police than Whites.

Research based on incident reports from deadly force encounters and laboratory-based experiments are the two major methods of research that have investigated the extent of racial and ethnic bias amongst the police. Studies based on incident reports are dependent on the accuracy and completeness of those reports. In addition, deadly force encounters tend to be so complex that it is difficult to reliably distinguish between race-bias effects and other threat markers such as suspect demeanor and behavior, based on recorded information. Experimental testing using "shoot/don't shoot" designs has greater internal validity than research based on incident reports because all other variables are constant; the only variation in stimulus prompts are suspect race and the presence or absence of a weapon. However, traditional experimental designs sacrifice external validity in order to control for the complex nature of actual police encounters. Critical external validity issues arise when generalizing from experimental results to racial bias in police use of deadly force in the field. Participants in traditional "shoot/don't shoot" experimental studies are seated in front of a desktop computer, they do not interact with the image in front of them, and they do not fire a weapon. The experimental designs used in prior laboratory studies bear almost no resemblance to an actual deadly force encounter.

The limitations from both research based on incident reports and research using traditional "shoot/don't shoot" experiments have created barriers to understanding the influence of suspect race and ethnicity on police use of deadly force. The goal of this research was to extend the base of evidence on racial and ethnic bias in police use of deadly force by improving study design. The current research used controlled randomized trials across three experiments, but was superior to previous laboratory experiments in that participants were placed in situations that more closely resembled real-life deadly force encounters. A set of realistic scenarios was developed for use in the deadly force judgment and decision-making simulators at the Washington State University (WSU) Simulated Hazardous Operational Tasks Laboratory; a laboratory dedicated to analyzing and understanding high-risk operational performance, including police use of deadly force. The laboratory simulators are the type used for training by many law enforcement agencies in the United States and around the world. The content of the scenarios was based on research examining incidents where officers were killed or assaulted during the past three decades (Federal Bureau of Investigation, Criminal Justice Information Services Division 2006). In addition, this is the first experimental study to go beyond the question of racial bias and investigate the possibility of ethnic bias against Hispanic suspects. Research based on incident reports has suggested that the police treat Hispanic suspects differently than they

treat White suspects (Liska and Yu 1992; Stone et al. 2010). This research is the first step in testing whether participants respond differently to White and Hispanic suspects in a controlled laboratory setting.

Literature review

Research based on incident reports from deadly force encounters and laboratory-based experiments are the two major methods of research that have been used to investigate the extent of racial and ethnic bias in deadly force judgment and decision making. Observational studies conducted by researchers in the field are not feasible given the rarity of deadly force encounters—less than 2 % of police–citizen encounters result in any use of force, and only a fraction of those develop into police shootings (Alpert and Dunham 2010). Research has, however, been conducted that observed police officers as they *trained* on deadly force judgment and decision-making simulators to assess the impact of suspect characteristics on decision making (Ho 1993). Ho observed that during training officers tended to be quicker to shoot armed Black suspects than armed White suspects, and more likely to fail to shoot armed White suspects than armed Black suspects (Ho 1993).

Although valuable to the research literature, observational studies have not focused on suspect race or ethnicity and lack the experimental control of laboratory-based studies. For example, Ho's study did not control for other factors that are thought to influence scenario difficulty, such as suspect demeanor or time pressure in the encounter. Furthermore, observation was conducted in a police training facility, making it impossible to control for the reactive effects of experimental arrangements (Campbell and Stanley 1963). When researchers observe police training they have very little control over the priming, stimuli, and feedback officers receive, or over the procedures officers are put through. This makes it difficult to differentiate racial or ethnic bias from other influences on officers' decisions to shoot, and limits the utility of research observing police during training as a research method for investigating the extent of racial and ethnic bias in deadly force judgment and decision making.

Prior research based on incident reports from deadly force encounters has investigated whether suspect race and ethnicity influence police use of deadly force. Much of this research has strongly supported the hypothesis that police officers are influenced by suspect race, independent of criminality. Takagi summarized this perspective with his statement that "the police have one trigger finger for whites and another for blacks" (Takagi 1974, 30). Sherman reported: "the demonstrably higher rates of police homicide for blacks strongly suggests racial discrimination on a national basis" (Sherman, 1979, 57). Research suggests that discrimination based on racial or ethnic cues is a primary cause of the disproportionate number of minority suspects shot by the police (Jacobs and O'Brien 1998; Sorenson et al. 1993). This hypothesis has been called the "quasi-labeling view" (Goldkamp 1976), and has also been called the "conflict approach" (Takagi 1974).

In 1992, Liska and Yu tested whether the percentage of racial and ethnic minority members in a city influenced police use of deadly force. They found that the higher the percentage of non-White community members in a city (including Blacks and Hispanics), the higher the rates of police use of deadly force (Liska and Yu 1992).

In a separate macro-level study, Sorenson and colleagues examined the FBI Supplemental Homicide Reports (SHR) from the largest cities in the United States and found that economic inequality and percent Black in a city significantly predicted increased numbers of shootings by the police (Sorenson et al. 1993). Similarly, in 1998, Jacobs and O'Brien investigated data from the FBI SHR and found that cities with a larger Black population, a recent growth in the Black population, and greater economic stratification based on race had significantly greater numbers of shootings by the police (Jacobs and O'Brien 1998). These studies based on incident reports support the notion that police officers *are* influenced by race or ethnicity when making decisions to shoot.

Other research based on incident reports takes a very different perspective. Some research has suggested that minority suspects, in particular Black suspects, pose a greater threat to the police, and that the influence of suspect race and ethnicity on police use of deadly force is insignificant when community-level violent crime rates and dangerousness of the offense underlying deadly encounters are controlled for (Inn et al. 1977; Fyfe 1978, 1982; Brown and Langan 2001; MacDonald et al. 2001). Inn et al. (1977) analyzed incident reports from a major metropolitan police department and found that officers shot significantly more Black suspects than White suspects relative to their numbers in the city's population; however, this discrepancy occurred at the criminal involvement level. Black suspects were more likely to shoot at the police than White suspects, relative to their numbers in the population (Inn et al. 1977). In fact, this study found that officers fired more shots at White suspects than at Black suspects, suggesting "perhaps, police behave more cautiously with Blacks because of departmental policy or public sentiment concerning treatment of Blacks" (Inn et al. 1977, 35).

In 1978, Fyfe found that 60 % of Black suspects shot by the police were carrying handguns compared to 35 % of White suspects (Fyfe 1978). In his 1982 study of police shootings in Memphis, Fyfe investigated Takagi's "trigger finger" assertion by analyzing situational determinants of police shootings. He felt that "to know whether police differentiate along racial lines with their trigger fingers, one must know something about the situations in which police shoot at members of different racial groups" (Fyfe 1982, 711). He found that situations in which police officers shot Black suspects were more life-threatening (determined by degree of officer injury) than situations involving White suspects.

In 1980, Meyer examined shooting incidents in Los Angeles and found that Blacks made up 18 % of the population, but were responsible for 44 % of attacks on officers (compared to Whites who made up 52 % of the population and 28 % of attacks on officers, or Hispanics, who made up 24 % of both the population and attacks on officers) (Meyer 1980). In a similar vein, Brown and Langan reported that between 1976 and 1998 Black suspects made up 12 % of the population but committed 43 % of felonious killings of officers, suggesting that the discrepancy in police use of deadly force against Black suspects may be due to Black suspects presenting a greater risk to the police (Brown and Langan 2001). In addition, in 2002, White found that situational predictors of deadly force incidents, and in particular relationships between these predictors, are important in determining causes of police shootings in Pennsylvania, and suggested that "the disproportionate percentage of black male shooting victims is at least partially a consequence of their involvement in gun

assaults against PPD officers… rather than discriminatory shooting practices involving white officers and black suspects" (White 2002, 746).

Research based on incident reports has several major limitations when used to examine the extent of racial and ethnic bias in deadly force judgment and decision making. First, this type of research depends on the accuracy and completeness of the information recorded. Although reports of suspect race and ethnicity are likely to be reasonably reliable, other aspects of an encounter may not. Studies suggest that perceptual memory can be severely distorted during deadly force encounters, and the ability to engage in rational thinking, which is required for comprehensive, narrative recall, is limited under conditions of trauma or high stress (Solomon and Horn 1986; Honig and Rolland 1998; Klinger 2004; Klinger and Brunson 2009). In a series of interviews with police officers following deadly encounters, Klinger found that diminished sound, tunnel vision, and the feeling of fast motion were commonly experienced (Klinger 2004; Klinger and Brunson 2009). The second major limitation in research based on incident reports is that the measures used are coarse and confounds are often not controlled for or even taken into account. Deadly force encounters tend to be complex, fast-paced and ambiguous. There are clear complications in extracting any one variable as causal out of a complex set of interactions that occur during police encounters. It is difficult to identify racial or ethnic bias as the cause of the disproportionate shooting of minorities when research based on incident reports cannot empirically measure this bias. In most cases, racial and ethnic bias cannot be parsed out unless other major variables are controlled for. During the last decade, research on racial bias in police use of deadly force has turned to a laboratory setting in an attempt to overcome this measurement problem.

Previous experimental research conducted in controlled laboratory settings has provided a way to directly assess the impact of suspect race on decision making by presenting participants with images of randomly paired Black and White suspects with weapons or neutral objects (such as wallets or cellular phones), to which they responded by pressing buttons labeled "shoot" or "don't shoot." Racial bias was inferred in two ways: (1) by whether participants were consistently quicker to shoot armed suspects of a particular race (measured by reaction time in seconds after the image is presented); and (2) if participants made "shooting errors" (pressing "shoot" for neutral images or pressing "don't shoot" for weapons) based on race. The intellectual basis for experimental studies of racial bias in deadly force judgment and decision making lies with a substantial body of research indicating that many Americans subconsciously associate Blacks with crime (e.g., Payne 2001; Payne et al. 2002; Amodio et al. 2003, 2004; Correll et al. 2006).

The results of repeated randomized trials suggest that civilian participants display implicit racial bias. For example, Payne (2001) found that, when participants were primed with images of Black faces on a computer screen, they were significantly more likely to automatically identify a "target picture" that followed as a weapon (even though half of the pictures were of tools). Furthermore, participants were quicker to pair pictures of weapons with Black faces, and tools with White faces. Research has also found that civilian participants were significantly more likely to mistake weapons for neutral objects when primed with White faces (Payne et al. 2002; Amodio et al. 2003, 2004). These results suggest that an implicit association

exists between Blacks and crime, without the awareness or intent of the participant (Payne 2001).

The first experimental study to examine racial bias in police officers was conducted by Eberhardt and colleagues in 2004. This study primed participants with abstract concepts, for example crime, basketball, arrest, violent, and shoot, then tested participants' attention orientation to Black and White faces in a series of images. Results showed that when officers were primed with the concept of crime, their attention was significantly oriented towards Black (more than White) faces (Eberhardt et al. 2004). Although these results are suggestive, their generalizability to racial bias in deadly force judgment and decision making is limited because these studies did not test decisions to shoot.

In 2005, Plant and colleagues conducted a series of experiments to determine whether the implicit racial bias previously found in civilian and police samples extended to racial bias in deadly force judgment and decision making (Plant et al. 2005). They presented a sample of college students with a series of images of either Black or White faces randomly superimposed with either an image of a weapon or an innocuous object such as a wallet. Participants had to respond by pressing a button marked "shoot" or one marked "don't shoot". Initial results showed a significant implicit association between Blacks and crime, similar to that reported by Payne (2001). Plant and Peruche also tested police officers using the same design (Plant and Peruche 2005), and found that, similar to the civilian participants, officers were more likely to shoot Black targets paired with neutral objects, and to fail to shoot White targets paired with guns. However, following repeated exposure to multiple trials where the race of the suspect was unrelated to the object, both civilian and police participants showed significant error reductions (Plant et al. 2005).

Correll and colleagues confirmed the results of prior experimental research using a more sophisticated research design over a series of studies. Participants were exposed to a series of video game simulations in which a still image slide show of armed and unarmed Black and White men appeared against a variety of backgrounds (an apartment building, a parking lot, a busy street, etc.). Similar to Plant and colleagues, participants were instructed to press a button labeled "shoot" or a button labeled "don't shoot", depending on whether the suspect was armed or unarmed. Civilian participants displayed significant racial bias in their reaction times and in their errors (Correll et al. 2007b, 2006). In 2007, this research was expanded to include police participants. Results showed that, similar to civilians, the police were significantly quicker to press "shoot" for armed Black suspects and quicker to press "don't shoot" for unarmed White suspects. However, unlike civilian participants, the sample of police officers showed no significant racial bias in their errors (they did not mistakenly shoot unarmed Black suspects or fail to shoot armed White suspects disproportionately). Correll and his colleagues suggested that: "by virtue of their training or expertise, officers may exert control over their behavior, possibly overriding the influence of racial stereotypes" (Correll et al. 2007a, 1014).

Previous experimental research has several weaknesses when used to examine the extent of racial and ethnic bias in deadly force judgment and decision making. The first limitation is that the test stimuli used in these experiments lack external validity, which raises the question of whether prior experimental results are generalizable to populations outside laboratory settings. The second limitation is the difference

between firing a weapon and pressing a button. Firing a handgun is a complicated endeavor; at minimum, it involves un-holstering, bringing the weapon to a ready position, aligning sights with the target, and ultimately pulling the trigger. Pushing a button is a simple reflex, dramatically different to the complex process involved in shooting a firearm. Furthermore, there is no active difference between pressing a "shoot" and a "don't shoot" button. The same action is required for a decision to shoot and a decision not to shoot, whereas in field encounters a decision not to shoot is marked by *inaction*. The third limitation is that prior experimental research has only tested participant responses to Black and White suspects. Thus, it can only increase our knowledge of racial bias in police use of deadly force. Research based on incident reports suggests that bias in police use of deadly force may not be limited to racial bias against Black suspects; ethnic bias against Hispanic suspects also appears to influence police use of deadly force. With the exception of the research reported here, no experimental research has tested for ethnic bias against Hispanic suspects.

Research goals and objectives

The limitations from both research based on incident reports and research using traditional "shoot/don't shoot" experiments have created barriers to understanding the influence of suspect race and ethnicity on deadly force judgment and decision making. The goal of this study was to extend the base of evidence on racial and ethnic bias in decisions to shoot by improving study design. The current research used experimental methods, but was superior to previous laboratory experiments in that participants were placed in situations that more closely resembled real-life deadly force encounters.

The first improvement in study design was greater external validity than previous experimental research with regard to the procedures participants were put through. Participants interacted with scenarios in a very immersive environment in which they fired real Glock handguns that had been modified for use in high-fidelity deadly force judgment and decision-making simulators. These laboratory simulators are the type used for training by many law enforcement agencies in the United States and around the world.

The second improvement in study design was greater external validity than previous experimental research with regard to the stimulus prompts participants received. A set of realistic scenarios was developed for use in the Washington State University (WSU) Simulated Hazardous Operational Tasks Laboratory; a large, highly controlled laboratory environment dedicated to analyzing and understanding high-risk operational performance, including police use of deadly force. The content of the scenarios was based on research examining incidents where officers were killed or assaulted during the past three decades (Federal Bureau of Investigation, Criminal Justice Information Services Division 2006).

The third improvement in study design was that White, Black and Hispanic suspects were presented in the scenarios (proportional to their involvement in police shootings), making this the first experimental study to go beyond the question of racial bias and investigate the possibility of ethnic bias against Hispanic suspects. Research based on incident reports has suggested that the police treat Hispanic

suspects differently than they treat White suspects (Liska and Yu 1992; Stone et al. 2010). This research is the first step toward testing whether participants respond differently to White and Hispanic suspects in a controlled laboratory setting.

This research design is more controlled than research based on incident reports, yet more sophisticated and face valid than experimental research using "shoot/don't shoot" button-pressing designs. By advancing the rigor of experimental designs in this arena, additional light is shed on the broad issue of the role that race and ethnicity play in the criminal justice system.

Research methods

Design

Three separate experiments ($n=102$) were conducted at the WSU Simulated Hazardous Operation Tasks laboratory, part of the Sleep and Performance Research Center in Spokane, to test participant decisions to shoot. Each of these experiments used a rigorous repeated measures design, where participants responded to multiple highly realistic and arousing scenarios in a deadly force judgment and decision-making simulator (see "Materials" below for details). From these three experiments, analyses were run on the impact of suspect race and ethnicity on participant decisions to shoot.

Participants

A total of 102 participants were tested in this study (see Table 1 for details). In the first experiment, 24 participants were recruited for a Defence Advanced Research Projects Agency (DARPA)-funded experiment that tested "expert" and "novice" performance in simulated deadly force encounters in high-fidelity simulators. Participants with at least 5 years of on-duty policing experience or participants who were active military with at least one tour of duty in combat infantry roles were recruited as the study's "experts." Participants with no police, military, or firearms experience were recruited as the study's "novices." Civilian participants were student and staff volunteers from the Spokane campus of WSU. Sample selection was non-random, participants volunteered for the study. In total, 12 experts (6 police and 6 military) and 12 novices participated. No power analysis was conducted because this experiment was a pilot study.

In the second experiment, 48 participants were recruited for a follow-on DARPA-funded experiment to test whether novices could be trained to the level of "expert" (based on the results of the first experiment) in simulated deadly force encounters in high-fidelity simulators. Again, these participants had no police, military, or firearms experience. These civilian participants were all recruited through local advertising, on a volunteer basis. A pilot study was conducted, based on the results of the first experiment, to determine sample size.

In the third experiment, 30 police participants were recruited for an in-house pilot study to examine police responses to various simulated operational tasks, including deadly force judgment and decision making. Participants were required to be active

Table 1 Characteristics of research participants, %/mean (range), (n=102)

		Experiment			Total
		1 (n=24)	2 (n=48)	3 (n=30)	
Race/ethnicity	White	83 %	85 %	100 %	89 %
	Black	4 %	2 %	0 %	2 %
	Hispanic	13 %	7 %	0 %	7 %
Sex	Male	75 %	46 %	99 %	73 %
	Female	25 %	44 %	1 %	27 %
Age (years)	All participants	30 (19–50)	26 (19–50)	37 (31–43)	31 (19–50)
Years of experience	Police	17 (9–24)	NA[a]	NR[b]	17 (9–24)
	Military	7 (6–10)	NA[a]	NR[b]	7 (6–10)
	Civilian	0	0	NA[c]	0

[a] *NA* not applicable, Experiment 2 tested only civilians

[b] *NR* not recorded, years of experience was not recorded for Experiment 3

[c] *NA* not applicable, Experiment 3 tested only police

duty patrol officers. Again, sampling was non-random—participants volunteered for the study, and no power analysis was conducted because this experiment was a pilot study for an upcoming Office of Navel Research (ONR)-funded study testing police performance on simulated operational tasks under various conditions.

Police participants were recruited from local urban and suburban police departments and Sheriff's offices, and military participants were recruited from local Army and Marine Corps units. Permission to recruit police officers and military personnel was obtained through direct contact with training offices. The training offices then informed police officers and military personnel to contact us if they were interested in participating. Word of mouth was employed to ensure that we received a sufficient number of volunteers.

All participants were required to be physically and psychologically healthy (established during screening—see below). There were no specific inclusion or exclusion criteria other than general health, and level of deadly force judgment and decision-making experience. Participants received monetary remuneration for taking part in the study. Participants were informed that their responses would remain confidential. The use of human subjects was approved by the WSU Institutional Review Board prior to participant recruitment. Given that the analysis on racial/ethnic bias was a secondary use of the data conceived of post hoc, participants were not provided with any reason to believe that they were being tested for racial or ethnic bias.

A breakdown of participant characteristics can be seen in Table 1. In the first experiment, the majority of participants were White males. The mean age of participants was 30 years (range: 19–50). For police and military participants, the mean number of years of experience was 13 (range: 6–24). In the second experiment, the mean age of participants was 26 (range: 19–50). Forty-four percent were female. The majority of participants were White. In the third experiment, the majority of participants were White males. The mean age of participants was 37 (range: 31–43).

Materials

These experiments were conducted at the WSU Simulated Hazardous Operational Tasks laboratory, which is equipped with two high definition deadly force judgment and decision-making simulators of the type commonly used in law enforcement training in the United States. Each simulator is fully enclosed in a sound deadened 7 × 5 m shooting range, with a 3.5 × 2 m screen at the far end on which HD video scenarios are displayed. The handguns used in these simulators are modified Glock model 21s, whose barrels have been replaced with infrared emitters that register shot placement on the screen and time of shot after threat exposure.

The scenarios used in this research were developed by two of the authors (_____and _____). In the spring of 2009, 60 scenarios requiring decisions about whether or not to use deadly force were filmed for use in the WSU Simulated Hazardous Operational Tasks Laboratory. The scenarios were filmed in naturalistic settings using paid professional actors recruited from a talent agency used by commercial film-makers in the Pacific North West. Suspects in the scenarios were Black, White, and Hispanic, and were either armed or unarmed (approximately 35 % of the time). Scenario content was based on data gathered from the last 30 years of incidents in which officers were killed or assaulted (Federal Bureau of Investigation, Criminal Justice Information Services Division 1992, 1997, 2006). These data were gathered and analyzed by A.J. Pinizzotto, E.F. Davis, and C.E. Miller III, and are known as the "deadly mix" research.

During scenario development, we used the collective deadly mix research to create scenarios that were as true to life as possible. The term "deadly mix" refers to the interactive combination of situation, officer, and offender, and how they converge. The deadly mix research reports that more than 80 % of deadly force encounters occur during either a disturbance call, an arrest situation/crime in progress, an investigation of suspicious persons/circumstances, or a traffic stop/pursuit. Approximately 30 % of deadly force encounters occur during the day, 40 % during the evening, and 30 % during the night. More than 80 % of suspects assault officers with a firearm (most frequently a handgun). The most common distance between suspect and officer at the time of assault is 5 ft. Approximately 50 % of offenders are White, and 50 % are non-White. More than 90 % of offenders are male, and more than 70 % are under the influence of alcohol or drugs at the time of the incident. These dynamics of deadly encounters were critical to the development of our scenarios. (Federal Bureau of Investigation, Criminal Justice Information Services Division 1992, 1997, 2006)

In order to analyze the impact of suspect race or ethnicity on participant decision making, we controlled and manipulated scenario difficulty. Scenario difficulty was theoretically grounded in Perrow's (1984) Normal Accident Theory (NAT) and Klinger's (2005) adaption of the theory to police encounters. NAT is based on the twin concepts of complexity (number of variables in an encounter) and coupling (how quickly change in one variable affects change in another). According to NAT, greater complexity and tighter coupling increase situational difficulty and the probability of an unintended or catastrophic outcome. Difficulty level in our scenarios was manipulated into three levels: naïve, intermediate, and journeyman. See Appendix A for sample naïve, intermediate, and journeyman scenario synopses. By controlling the

variables that are thought to influence situational difficulty (number of people in the encounter, speed with which the scenario unfolds, suspect demeanor, suspect intoxication, and deceptive behavior of suspect), we were able to assess the impact of suspect race and ethnicity on participant decisions to shoot.

Procedures

Participants were screened for suitability approximately 1 week prior to experiments. They were required to be physically and psychologically healthy, i.e., no clinical disorders and/or illnesses (by history, physical examination, and questionnaires). Prior to screening, participants were briefed about the study process, risks, and potential benefits of participating in the study, then asked to sign informed consent before participating in the study. Note that the issue of scenario suspect race/ethnicity was *not* raised with participants during the consent process because the goal of the experiments was to assess unrelated deadly force decision-making issues.

On the day of the experiments, participants were outfitted with belts and holsters, taught how to operate the weapon system, and given an orientation regarding range procedures and the following rules of engagement:

> The goal of a police officer in a deadly force encounter is to accurately identify a threat and neutralize it, while minimizing harm to bystanders, officers, and suspects.

During the day-long experiments, each participant responded to between 10 and 27 scenarios in a deadly force judgment and decision-making simulator (AIS PiSim, Seattle, WA, USA). Participants in the first experiment responded to 27 scenarios, participants in the second experiment responded to 18 scenarios, and participants in the third experiment responded to 10 scenarios. These scenarios were randomly drawn from the pool of 60, then randomized and organized into sets for each participant to avoid learning effects.

Participants in the second experiment (n=48 civilians) received a pre-test training session in the simulators where they each practiced a scenario and received feedback from local law enforcement trainers (this was because the primary purpose of that particular experiment was to see if novices could be trained to expert standards using simulated encounters).

All participants had a 3-min rest period between each scenario, during which they sat quietly in a chair in a corner of the range, and a 30-min rest after completing a set of scenarios. Participants were monitored to assure that they did not discuss the scenarios or the simulators with other research participants.

Study variables and analytical strategy

In each of these three experiments, participants responded to multiple scenarios. This resulted in two levels of variation for the response variables—between-participant variance, and within-participant variance based on repeated observations (scenario-driven variance). The study variables of interest in the analyses presented here were scenario-level variables: race/ethnicity of suspect (White, Black, Hispanic), and

scenario difficulty level (naïve, intermediate, journeyman). In keeping with prior experimental research, the dependent response variables were reaction time to shoot, and two dichotomous variables: shooting an *unarmed* suspect (error A), and failing to shoot an *armed* suspect (error B).

In the instance of failing to shoot an armed suspect, reaction time was automatically coded as non-applicable to avoid skewing reaction time data. For each study variable, dummy variables were created to assess the specific influence of each race/ethnicity category (using White suspects as the reference group) and difficulty level (using naïve scenarios as the reference group) on reaction time, error A, and error B. Each dependent response variable was analyzed independently.

Multi-level regression models were used to examine factors associated with variation in response variables. This analytical strategy was chosen because each participant responded to multiple scenarios, requiring multi-level mixed effects models to account for the lack of independence among observations. In order to determine whether the assumption of independence was violated, unconditional means models were conducted for each dependent response variable. When clustering of observations around participants occurred (violating the assumption of independence), R project for statistical computing software was used to create multi-level regression models. Where no clustering was observed, multi-variate and binary logistic regression models were used to analyze the data in SPSS. By separating between-participant variation from within-participant variation, this analytical technique reduced the likelihood that error related to individual differences influenced results, and increased statistical power.

Research findings

Tests for normality

An examination of the data showed that across the three experiments reaction time was positively skewed in scenarios with White, Black, and Hispanic suspects, and across each level of difficulty. The Kolmogrov–Smirnov test was significant for each category. However, the outliers that were identified represented valid data points, specific to scenarios. For this reason outliers were not removed or transformed. Furthermore, we did not wish to transform the data as it would have changed the unit of measurement from seconds, which is easily interpreted. Because outliers existed for scenarios with White, Black, and Hispanic scenarios, the analysis of racial and ethnic bias was not affected by a lack of normality.

Variation between participants

Unconditional means models run on each response variable across each of the three experiments showed that observations were *not* clustered around participants. The intra-class correlation coefficients (ICC) for reaction time, error A, and error B were all smaller than 0.03, indicating that nearly all of the variation in response variables occurred within participants (>97 %). This was strong evidence that the assumption

of independence of observations was *not* violated. Because of this, multi-variate and binary logistic regression techniques were the most suitable.

Reaction time

For each experiment, multi-variate regression models were run to determine the impact of each study variable on reaction time. In the first experiment (n=12 civilian, 6 active duty police, and 6 active duty military participants), the overall regression model was significant (f=19.41; df=6, 400; p=< .001) and predicted 23 % of the variance in reaction time (r^2=0.23).

Table 2 shows that participants took significantly (0.80 s) longer to shoot Black suspects than White suspects (t=3.25; p<.001). There was no significant difference in reaction time between shooting Hispanic suspects and White suspects. Participants took significantly longer to shoot in journeyman (t=7.83; p<.001) and intermediate (t=2.83; p<.005) scenarios than they took to shoot in naïve scenarios (2.26 s and 0.81 s longer, respectively). A significant interaction existed between suspect race/ethnicity and scenario difficulty (t=5.90; p<.001); participants took longest to shoot Black suspects in journeyman scenarios.

In the second experiment (n=48 civilian participants), the overall regression model was significant (f=12.25; df=4, 563; p<.001) and predicted 8 % of the variance in reaction time (r^2=0.08).

Table 3 shows that participants took significantly (0.68 s) longer to shoot Black suspects than White suspects (t=3.85; p<.001). There was no significant difference in reaction time between shooting Hispanic suspects and White suspects. Participants took significantly longer to shoot in journeyman (t=4.23; p<.001) and intermediate (t=6.75; p<.001) scenarios than they took to shoot in naïve scenarios (1.12 and 1.52 s longer, respectively). A significant interaction existed between suspect race/ethnicity and scenario difficulty (t=2.00; p<.05); participants took longest to shoot Black suspects in intermediate and journeyman scenarios.

In the third experiment (n=30 active duty police participants), the overall regression model was significant (f=10.68; df=4, 156; p<.001) and predicted 22 % of the variance in reaction time (r^2=0.22).

Table 4 shows that participants took significantly (1.34 s) longer to shoot Black suspects than White suspects (t=4.14; p<.001). There was no significant difference in reaction time between shooting Hispanic suspects and White suspects. Participants took significantly (0.93 s) longer to shoot in journeyman scenarios than they took to

Table 2 Summary of regression analysis for variables predicting reaction time to shoot in experiment 1 (n=24)

Variable	B	SE B	β
Black suspects	0.80	0.25	0.15**
Hispanic suspects	0.07	0.46	0.01
Police and military participants	−0.30	0.24	−0.06
Intermediate scenarios	0.81	0.29	0.14*
Journeyman scenarios	2.26	0.29	0.39**

*p<.005, ** p<.001

Table 3 Summary of regression analysis for variables predicting reaction time to shoot in experiment 2 (n=48)

Variable	B	SE B	β
Black suspects	0.68	0.18	0.17**
Hispanic suspects	0.17	0.71	0.01
Intermediate scenarios	1.52	0.22	0.39**
Journeyman scenarios	1.12	0.26	0.24**

**p<.001

shoot in intermediate scenarios (t=4.23; p<.001) (participants in this study did not respond to naïve scenarios). A significant interaction existed between suspect race/ethnicity and scenario difficulty; participants took the longest to shoot Black suspects in journeyman scenarios (t=29.63; p<.001).

Shooting an unarmed suspect

For the first and third experiments, binary logistic regression models were run with error A (shooting an unarmed suspect) as the dependent response variable (too few errors were made by participants in the second experiment to warrant analysis—see "Discussion" for implications).

In the first experiment (n=12 civilian, 6 active duty police, and 6 active duty military participants), the regression model was significant (χ^2=27.10; df=5; p<.001) and explained 18 % of the variance in likelihood of shooting an unarmed suspect (r^2=0.18). Participants were less likely to shoot unarmed Black suspects than unarmed White suspects (p<.005). Using the Inverse Odds Ratio (dividing the Odds Ratio into 1), we calculated that participants were almost *six times* less likely to shoot unarmed Black suspects than they were to shoot unarmed White suspects (Inverse Odds Ratio = 1/0.17=5.88). There was no significant difference between the likelihood of shooting unarmed Hispanic suspects and unarmed White suspects. Participants were almost ten times more likely to shoot unarmed suspects in journeyman (p<.005) and intermediate (p<.005) scenarios than they were in naïve scenarios. There was no interaction between suspect race/ethnicity and scenario difficulty for shooting an unarmed suspect.

In the third experiment (n=30 active duty police participants), the regression model was significant (χ^2=31.77; df=3; p<.001) and predicted 23 % of the variance in likelihood of shooting an unarmed suspect (r^2=0.23). Participants were less likely to shoot unarmed Black suspects than unarmed White suspects (p<.005). Using the Inverse Odds Ratio we calculated that participants were *25 times* less likely to shoot unarmed Black suspects than they were to shoot unarmed White suspects

Table 4 Summary of regression analysis for variables predicting reaction time to shoot in experiment 3 (n=30)

Variable	B	SE B	β
Black suspects	1.34	0.32	0.31**
Hispanic suspects	0.05	0.54	0.01
Journeyman scenarios	0.93	0.32	0.21**

**p<.001

(Inverse Odds Ratio = 1/0.04=25.00). There was no significant difference between the likelihood of shooting unarmed Hispanic suspects and unarmed White suspects. Scenario difficulty did not significantly predict shooting an unarmed suspect. There was, however, a significant interaction between suspect race/ethnicity and scenario difficulty; participants were most likely to shoot unarmed White suspects in journeyman scenarios ($p<.05$).

Failing to shoot an armed suspect

For the first and third experiments, binary logistic regression models were run with error B (failing to shoot an armed suspect) as the dependent response variable (too few errors were made by participants in the second experiment to warrant analysis—see "Discussion" for implications).

In the first experiment ($n=12$ civilian, 6 active duty police, and 6 active duty military participants), the regression model was significant ($\chi^2=30.64$; $df=6$; $p<.001$) and explained 18 % of the variance in likelihood of failing to shoot an armed suspect ($r^2=0.18$). Participants were *five times* more likely to fail to shoot armed Black suspects than armed White suspects ($p<.005$). There was no significant difference between the likelihood of failing to shoot armed Hispanic suspects and armed White suspects. Participants were slightly more than five times as likely to fail to shoot armed suspects in journeyman scenarios ($p<.005$) than they were in naïve scenarios. There was no significant difference between failing to shoot armed suspects in intermediate and naïve scenarios. A significant interaction existed between suspect race/ethnicity and scenario difficulty ($t=3.37$, $df=177$; $p<.001$). Participants were most likely to fail to shoot armed Black suspects at the highest level of difficulty.

In the third experiment ($n=30$ active duty police participants), the regression model was insignificant—participants were equally likely to fail to shoot armed White, Black, and Hispanic suspects at each level of difficulty.

Discussion

Summary of results

These three experiments tested police, military, and civilian participants in a deadly force judgment and decision-making simulator. Subsequent analyses revealed that racial and ethnic bias appeared to exist, but in the opposite direction than would be expected from prior experimental studies (Plant and Peruche 2005; Correll et al. 2007a). Across each of the three experiments, police, military, and civilian participants were significantly *slower* to shoot Black suspects than Whites or Hispanics. In addition, in two of the experiments, participants were significantly *more likely* to shoot unarmed White suspects than Blacks or Hispanics. Finally, in one of the experiments, participants were significantly *more likely* to fail to shoot armed Black suspects than Whites or Hispanics. In sum, the participants in this experiment showed significant evidence of bias *favoring* Black suspects, rather than discriminating against them.

Review of the results in light of prior research

The results from these experiments are markedly different to the results of prior experimental research. Across experiments, participants took significantly longer to shoot Black suspects than White or Hispanic suspects, particularly in the most difficult scenarios. It is critical to note that, although our research found that police, military, and civilian participants had similar reaction times to shoot and made similar shooting errors, this does not mean that police and military participants did not outperform civilian participants. The police and military participants had better shooting accuracy, fired faster follow-on shots, were far more interactive with the scenarios (for example, shouting at suspects: "drop your weapon or I will shoot!"), and had superior command presence than civilian participants. None of these variables were considered to have relevance to this analysis, which is why they were not included. It is important not to interpret our findings to mean that civilians are as successful at deadly force judgment and decision making as police and military participants—simply that they tend to respond to suspect race and ethnicity in the same way. In addition, the lack of shooting errors in the second experiment (which had 48 civilian participants with no police, military, or shooting experience) was likely due to the intensive training session they received immediately before testing. This is a more reasonable explanation for their lack of errors than them "outperforming" officers—given that the 12 civilians who did not receive pre-test training made errors in the first experiment.

Our primary finding that participants were more hesitant to shoot Black suspects than White or Hispanic suspects is in direct contrast to prior experimental findings that participants are significantly quicker to shoot Black suspects (Correll et al. 2007a). In two of the experiments, participants were significantly more likely to shoot unarmed White suspects, and in one of the experiments, participants were more likely to fail to shoot armed Black suspects. These findings are also in direct contrast to Plant and colleagues who found that participants were more likely to shoot unarmed Black suspects and fail to shoot armed White suspects, with a lessening of racial bias after repeated randomized trials (Plant et al. 2005). Why would our novel design that better represents the actual conditions police officers are exposed to produce such strong effects in the unexpected direction? It is possible that the differences between our results and prior experimental results reflect our subjects' own personal biases, moral concerns, cultural awareness, societal pressures, administrative effects, or desire to appear unbiased.

Implications and competing explanations

Given the limited diversity and non-random sampling of research participants, research implications are tentative. However, our main finding that participants were more hesitant to shoot Black suspects than Whites or Hispanics was replicated across three separate experiments, and is in direct contrast to previous experimental findings. As such, we believe that this finding deserves investigation. Several competing explanations are possible.

Our findings may indicate a lack of implicit (or subconscious) racial bias connecting Black suspects and dangerousness; perhaps indicative of a growing cultural

awareness that Blacks and crime are not synonymous. We, however, find this explanation unlikely. If Blacks and crime where not related in the minds of our subjects, it would stand to reason that they would be treated in the same way as White and Hispanic suspects. This was not the case. If implicit racial bias is the explanation, it would appear that subjects feel Black suspects are less dangerous than Whites or Hispanics; that they are biased in favor of Black suspects. Although possible, we feel that this is unlikely. We are currently investigating the neurophysiological responses of participants to see if they display more or less of a "threat" response (sympathetic arousal) to Black suspects. If participants do display a greater threat response to Black suspects, indicative of a subconscious association between Black suspects and dangerousness, but still are more hesitant to shoot Black suspects, it will provide further evidence that some conscious cognitive process is actively guarding against bias. We also plan on testing participants using both the Implicit Association Test (IAT) and our scenarios, to investigate differences between implicit and behavioral bias. The results of this upcoming research will tell us whether subjects lack subconscious racial bias, or whether something else is making them more hesitant to shoot Black suspects.

It is possible that the results may have no relevance to subconscious racial bias, but imply a behavioral "counter-bias" due to real-world concern over discipline, liability, or public disapproval. This potential explanation has particular relevance for the study's police participants (n=36), for whom an "administrative effect" may influence decisions to shoot. Many law enforcement agencies have tightened their formal policies and increased cultural awareness training, due to public pressure and awareness that the consequences of deadly force encounters tend to be graver when the suspect in question is a member of a racial or ethnic minority (White 2001). The law enforcement agencies from which our participants were sampled offer cultural awareness training, conduct in-depth evaluation of shooting incidents (the frequency of which are representative of a large, metropolitan city), and provide discipline for any misuse of force. Although an administrative effect may have influenced police participants, it could not have affected civilians, who have not been exposed to police training, evaluation, or discipline. Given that civilians and police participants responded in similar ways, it is unlikely that an administrative effect is solely responsible for these results. Nonetheless, it is possible that moral concerns or societal pressures to guard against racism may have influenced participant responses. Perhaps public concern over the treatment of racial and ethnic minority suspects, and consequent pressure on the police to guard against bias, are working.

It is also possible that the results may indicate self-awareness or a desire to appear unbiased during testing. Despite participants not being made aware that they were being tested for racial or ethnic bias, it is arguable that police officers are sensitive to concerns about racial bias. As such, they may have suspected that their responses would have been monitored for biases, and tailored their behavior accordingly. Again, this explanation does not fit particularly well with civilian participants who have not been exposed to the same societal pressure, making it unlikely that self-awareness was the primary cause of the results. On this point, it is also important to note that, although the proportion of Black, White, and Hispanic suspects in the scenarios represented police shootings nation-wide, Black suspects were slightly over-represented based on the demographics of police shootings in the geographical

areafrom which participants were sampled. This may also have made participants aware that suspect race may be considered in subsequent analysis. The fewer number of Hispanic suspects than Black or White suspects may also have resulted in insufficient power to differentiate between participant responses to White and Hispanic suspects. Given that there are no other experimental data on bias against Hispanic suspects, this apparent lack of bias against Hispanics needs further investigation.

Finally, it is possible that other explanations exist, such as the political climate in the area at the time of experimentation. We feel this is unlikely, because there were no racially charged events or news stories in the area at the time.

Limitations and future research

The current study has a number of limitations. First, as with any laboratory-based study, the artificial nature of the research must be taken into account. By presenting participants with repeated simulations of deadly encounters, we likely generate an artificially high sensitivity to danger in participants. In the real world, encounters such as these are rare, so the condensed nature in which these simulations are presented is not externally valid, despite scenario realism. However, the very fact that these encounters are rare and not easy to study ethnographically, combined with the large-scale attention they attract, provide ample justification for investigating their causes in a laboratory environment.

Second, the scenarios used in this research are realistic and complex, which increases external validity, but also tends to decrease experimental control compared to other experimental studies in which stimulus prompts are *exactly* the same except for suspect race. For example, it is possible that variation in actors' performances in the scenarios may have created variation in participant decisions during testing that was attributed to suspect race or ethnicity. We do not think this is the case because of the professional actors' skilled performances. It is important to note that actors that appeared in more than one scenario were not consistently armed, or consistently unarmed. This was to avoid the possibility of participants recognizing a particular actor as a "bad guy" or a "good guy."

Third, generalizability from these three experiments should not be assumed. Only six military participants were tested, excluding any generalizability of these results to military populations. In addition, our police samples over-represent White males, making the generalizability of our results to law enforcement populations tentative. We do not imply that these findings can be generalized to the U.S. law enforcement population. However, our samples are representative of Eastern Washington law enforcement (which is dominated by White males). Replication outside Eastern Washington is necessary to see whether our significant findings generalize to more diverse samples. Given the voluntary nature of participation, selection bias may also have influenced experimental results. Despite these limitations, our results were replicated across three separate samples, providing support for the reliability of our research findings.

Forth, measuring racial and ethnic bias was *not* the primary research goal of the experiments. The scenarios were not matched by race/ethnicity but were randomly assigned. It is possible that a more careful selection of scenarios may have provided a

more precise measurement of racial and ethnic bias. The major benefit of secondary use of the data was that neither the participants nor the researchers were made aware that their responses were going to be analyzed for racial or ethnic bias, reducing the likelihood that participant desire to appear unbiased influenced the results.

Fifth, participant characteristics were not analyzed for their impact on racial and ethnic bias in decisions to shoot. Participant age, gender, race, ethnicity, years of combat experience, and prior deadly force experience may have influenced decisions to shoot. There was not enough variation in participant race or ethnicity to analyze its effect on performance. In particular, it is possible that police participants who had prior deadly force encounters with suspects of a particular race or ethnicity may have been influenced during experimentation. Unfortunately, this information was not gathered at the time of experimentation. Having said that, the vast majority of variation in participant responses was within-participant or scenario-driven, implying that between-participant differences (such as age, gender, race, ethnicity, and years of combat experience) had no impact on reaction time to shoot or shooting errors.

Sixth, reaction time to shoot and shooting an unarmed suspect are likely related—additional time taken before making the decision to shoot or not shoot will likely influence whether or not the participant has mistakenly shot an unarmed suspect. More time to deliberate over a decision will typically reduce the likelihood that that decision is an erroneous one. In tightly controlled, button-pressing "shoot/don't shoot" experiments such as Correll's (2006), it is possible to fully distinguish between reaction time and shooting errors. However, in the real world, the relationship between response time to shoot and shooting errors is harder to disentangle. As such, we feel that presenting our major research findings as separate but related is empirically valid.

Finally, although the scenarios used in this research far better represent officer-involved shootings than still images of armed or unarmed suspects, they are still "shoot/don't shoot" stimulus prompts. As such their accurate representation of a police–citizen encounter is constrained. They are not interactive, less than lethal force is not an option when responding to them, and their outcome is determined entirely by decisions to shoot or not to shoot. They are accurate snippets of deadly encounters, but they are limited in that they are predetermined, and cannot be manipulated based on participant performance in the build-up to a shooting. To address this limitation, we have filmed a new set of scenarios that are equally engaging and arousing, but are also interactive. Our new scenarios have multiple branching options that enable the researcher to manipulate the outcome of the scenario based on how the participant is managing the encounter. For example, if the participant uses social interaction skills in an attempt to de-escalate a volatile encounter, the scenario can branch to avoid a deadly encounter. Alternatively, if a participant makes no effort to talk to the suspect, the scenario can branch to result in use of deadly force. These new scenarios are a step closer to the realistic simulation of potential officer-involved shootings. Using these scenarios, we will be able to explore the extent to which racial and ethnic differences tend to influence the *behaviors* of officers during police–citizen encounters in a manner that increases the likelihood that deadly force will be used.

Conclusions

The goal of this study was to analyze the influence of suspect race and ethnicity on decisions to shoot in deadly force judgment and decision-making simulators using realistic scenarios as stimulus prompts. In contrast to prior research, we found evidence of bias *favoring* Black suspects across police, military, and civilian samples. These results were replicated across three separate experiments. Participants were *slowest* to shoot Black suspects; and when errors were made, they were *less likely* to shoot unarmed Black suspects than unarmed Whites or Hispanics, and tended to be *more likely* to fail to shoot armed Black suspects than armed Whites or Hispanics. There were no significant differences in participant decisions to shoot White and Hispanic suspects.

By advancing the rigor of experimental designs, additional light is shed on the broad issue of the role that race and ethnicity play in the criminal justice system. Using deadly force judgment and decision-making simulators with engaging, arousing, and realistic scenarios as a platform to measure racial and ethnic bias is both more controlled than research based on incident reports, and more externally valid than experimental research using "shoot/don't shoot" button-pressing designs. This novel research design provides a bridge between highly controlled experimental studies and research based on reports of real-life encounters. By using an externally valid research design, our findings challenge the results of less robust experimental designs. Nearly 40 years ago, Takagi boldly stated that "the police have one trigger finger for whites and another for blacks" (Takagi 1974, 30). The results of the three experiments presented here suggest that this is true, though not in the way that Takagi intended. *For the participants in these experiments the "trigger finger" for Black suspects was the more hesitant*. Future research should explore the generalizability of these findings, determine whether bias favoring Black suspects is a consequence of administrative measures (e.g., education, training, policies, and laws), and identify the cognitive processes that underlie this phenomenon.

Appendix A

Name: SHOT_Veh_01_N
Type: Vehicle stop
Difficulty: Naive
Synopsis: The vehicle contains a Black male driver, no passengers. The suspect appears intoxicated; his movements are sluggish. He appears to search for his license then slowly pulls out a handgun, points it at the officer and opens fire.

Name: SHOT_SusP_04_N
Type: Investigation of suspicious persons
Difficulty: Naive
Synopsis: A White male suspect in an empty parking lot appears to be breaking into a car with a knife. He complies with the officer commands; he puts the knife on the roof of the car. He then reaches behind his back and pulls out a wallet.

Name: SHOT_SusP_20_J
Type: Investigation of suspicious persons
Difficulty: Journeyman
Synopsis: A Black male, a White female and a White male appear to be dealing drugs on the roof of an empty warehouse. The Black male and White female take off running and the White male quickly pulls out a handgun, points it at the officer and opens fire.

Name: SHOT_DwelH_08_I
Type: Domestic disturbance
Difficulty: Intermediate
Synopsis: A White male has hold of a White female. They are both shouting and swearing. The male drags the female down a hallway and pulls out a handgun, shoots the female, then turns on the officer and opens fire.

Name: SHOT_DwelH_14_J
Type: Domestic disturbance
Difficulty: Journeyman
Synopsis: A Black male and a Black female who is holding a baby are shouting at each other. They are both at the bottom of a flight of stairs going down to the basement. The male pulls out a handgun, points it at the officer and opens fire

References

Alpert, G., & Dunham, R. (2010). Policy and training recommendations related to police use of CEDs: overview of findings from a comprehensive national study. *Police Quarterly, 13*(3), 235–259.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84*(4), 738–753.

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 unintentional race bias. *Psychological Science, 15*(2), 88–93.

Brown, J., & Langan, P. (2001). *Policing and Homicide, 1976–98: Justifiable Homicide by Police, Police Officers Murdered by Felons. U.S. Department of Justice*. Washington: Bureau of Justice Statistics.

Campbell, D., & Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.

Cooper, M. (February 5, 1999). "Officers in Bronx Fire 41 Shots, And an Unarmed Man Is Killed." The New York Times.

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.

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007a). Across the thin blue line: police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology, 92*(6), 1006–1023.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007b). Stereotypes and racial bias in the decision to shoot. *European Journal of Social Psychology, 37*(6), 1102–1117.

Deakin, T. (1988). Police Professionalism: The Renaissance of American Law Enforcement. Springfield, Ill.; Charles C. Thomas.

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.

Federal Bureau of Investigation, Criminal Justice Information Services Division. (2006). *Violent Encounters: A Study of Felonious Assaults on Our Nation's Law Enforcement Officers*. Washington: U.S. Department of Justice.

Fyfe, J. J. (1978). *Shots Fired: An Analysis of New York City Police Firearms Discharges*. PhD dissertation, State University of New York at Albany.

Fyfe, J. J. (1982). Blind justice: police shootings in Memphis. *The Journal of Criminal Law and Criminology, 73*, 702–722.

Fyfe, J. (2000). Urban Policing in Australia and the United States, In Policing Issues, Issue #1, pp. 1–12

Geller, W. A. (1982). Deadly force: what we know. *Journal of Police Science and Administration, 10*, 151–177.

Goldkamp, J. S. (1976). Minorities as victims of police shootings: interpretations of racial disproportionality and police use of deadly force.". *Justice System Journal, 2*, 169–183.

Ho, T. (1993). Individual and situational determinants of the use of deadly force: a simulation. *American Journal of Criminal Justice, 18*(1), 41–60.

Honig, A. & Roland, J. (1998). "Shots Fired: Officer Involved." The Police Chief.

Inn, A., Wheeler, A., & Sparling, C. (1977). The effects of suspect race and situation hazard on police officer shooting behavior. *Journal of Applied Social Psychology, 7*(1), 27–37.

Jacobs, D., & O'Brien, R. (1998). The determinants of deadly force: a structural analysis of police violence. *The American Journal of Sociology, 103*(4), 837–862.

James, G. R., & Pasquale-Styles, M. (2009). Firearm Deaths by Law Enforcement. *Journal of Forensic Sciences, 54*(1), 185–188 (Blackwell Publishing Limited).

Klinger, D. (2004). *Into the Kill Zone: A Cop's Eye View of Deadly Force*. San Francisco: Jossey-Bass.

Klinger, D. (2005). *Social Theory and the Street Cop: The Case of Deadly Force. Ideas in American Policing, No. 7*. Washington, D.C: Police Foundation.

Klinger, D., & Brunson, R. (2009). Police officers perceptual distortions during lethal force situations: informing the reasonableness standard. *Criminology and Public Policy, 8*(1), 117–140.

Kobler, A. (1975). Figures (and perhaps some facts) on police killing of civilians in the United States 1965–1969. *Journal of Social Issues, 31*(1), 185–191.

Liska, A., & Yu, J. (1992). Specifying and Testing the Threat Hypothesis: Police Use of Deadly Force. In A. E. Liska (Ed.), *Social Threat and Social Control* (pp. 53–68). Albany: State University of New York Press.

Lopez, R.J. "No evidence Ruben Salazar was targeted in killing, report says." Los Angeles Times, February 19, 2011.

MacDonald, J., Kaminski, R., Alpert, G., & Tennenbaum, A. (2001). The temporal relationship between police killings of civilians and criminal homicide: a refined version of the danger-perception theory. *Crime & Delinquency, 47*, 155–172.

Meyer, M. (1980). Police shootings of minorities: the case of Los Angeles. *The Annals of the American Academy of Political and Social Science, 452*, 98–110.

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*, 1–12.

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

Perrow, C. (1984). *Normal Accidents: Living With High Risk Technologies*. Princeton, NJ: Princeton University Press.

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 Social Psychology, 41*, 141–156.

Robin, G. D. (1963). Justifiable homicide by police officers. *Journal of Criminal Law, Criminology, and Police Science, 54*, 225–229.

Sherman, L., & Langworthy, R. (1979). Measuring Homicide by Police Officers. *The Journal of Criminal Law and Criminology, 70*, 546–560.

Solomon, R. M., & Horn, J. M. (1986). Post-Shooting Traumatic Reactions: A Pilot Study. In J. T. Reese & H. A. Goldstein (Eds.), *Psychological Services for Law Enforcement*. Washington, D.C.: U.S. Government Printing Office.

Sorenson, J., Marquart, J., & Brock, D. (1993). Factors relating to killings of felons by police officers: a test of the community violence and conflict hypotheses. *Justice Quarterly, 10*(3), 417–440.

Stone, C., Carter, Z., Belfiore, T., Bully-Cummings, E., Daughtry, H., Farrell, M., Gascon, G., Rice, A., Rice, L. & Hewitt, D. (2010). Reducing Inherent Danger: Report of the Task Force on Police-on-police Shootings. New York State Task Force on Police-on-Police Shootings.

Takagi, P. (1974). A garrison state in a "democratic" society. *Crime and Scholarly Justice, 1*, 27–33.

U.S. Department of Justice. (2001). *Policing and homicide, 1976–98: Justifiable homicide by police, police officers murdered by felons (NCJ180987)*. Washington, DC: Bureau of Justice Statistics.

U.S. Department of Justice, Federal Bureau of Investigation, Killed in the Line of Duty: A Study of Selected Felonious Killings of Law Enforcement Officers (Washington, DC, 1992).

U.S. Department of Justice, Federal Bureau of Investigation, In the Line of Fire: Violence Against Law Enforcement (Washington, DC, 1997).

US-House-Committee-on-the-Judiciary-Subcommittee-on-Criminal-Justice. (1989). *Police Use of Deadly Force. Hearing…May 8, 1987*. Washington, DC: U.S. Government Printing Office.

Walker, S., Spohn, C., DeLone, M. (2004). *The Color of Justice: Race, Ethnicity and Crime in America*. 3rd edn. Belmont, CA: Wadsworth/Thompson

White, M. D. (2001). Controlling police decisions to use deadly force: reexamining the importance of administrative police. *Crime & Delinquency, 47*(1), 131–151.

White, M. D. (2002). Identifying situational predictors of police shootings using multivariate analysis. *Policing: An International Journal of Police Strategies and Management, 25*(4), 726–751.

Lois James Ph.D., is a research assistant professor of criminal justice and criminology at WSU Spokane. Dr. James works in the Sleep and Performance Research Center and is currently Co-Principle Investigator on $2.9 million in externally funded research projects. Dr. James studied psychology at Trinity College Dublin, and received her Ph.D. in criminal justice from WSU in 2011.

Bryan Vila Ph.D., is a professor of criminal justice and criminology at WSU Spokane. Prior to joining WSU in July 2005, he directed the Division of Crime Control and Prevention Research at the U.S. Department of Justice's National Institute of Justice. Dr. Vila has served as Principal Investigator or Co-Principal Investigator on $7.6 million in externally funded research projects. He has brought eight grants/contracts totaling $5.5 million to WSU in the past six years. Dr. Vila received his Ph.D. in Ecology from the University of California, Davis in 1990.

Kenn Daratha Ph.D., serves on the faculty at the WSU College of Nursing and at the University of Washington School of Medicine. He teaches doctoral statistics, evidence-based practice, and clinical epidemiology. His research interests include improving health outcomes for hospitalized patients. His most recent publication in the Clinical Journal of the American Society of Nephrology was titled 'Risks of Subsequent Hospitalization and Death in Patients with Kidney Disease'. His manuscript titled 'Co-Occurring Mood Disorders among Hospitalized Patients and Risk for Subsequent Medical Hospitalization' was just accepted by the journal General Hospital Psychiatry.