JUSTICE QUARTERLY
https://doi.org/10.1080/07418825.2020.1760331

The effect of suspect race on police officers' decisions to draw their weapons

John L. Worrall[a], Stephen A. Bishopp[b] and William Terrill[c]

[a]Criminology, University of Texas at Dallas, Richardson, TX, USA; [b]Dallas Police Department, University of Texas School of Public Health at Houston, Dallas, TX, USA; [c]Watts College of Public Service and Community Solutions, School of Criminology and Criminal Justice, Arizona State University, Phoenix, United States

ABSTRACT
Researchers are working to identify appropriate benchmarks for exploring racial bias in the officer-involved shooting (OIS) context. Two recent studies benchmarked OIS against incidents in which officers drew weapons but did not shoot. A problem is that the decision to draw a weapon may itself be subject to bias. Using 2017 use-of-force data from the Dallas Police Department, we modeled officers' decisions to draw their weapons as a function of suspect race and other suspect, officer, and incident characteristics. We benchmarked by limiting analyses to arrest and active aggression cases, thereby excluding interactions in which it was less likely suspects would have had weapons drawn against them. The key finding was that black suspects were no more or less likely to have weapons drawn against them than other suspects.

ARTICLE HISTORY
Received 17 October 2019
Accepted 20 April 2020

KEYWORDS
Officer-involved shootings; use of force; deadly force; suspect race

Introduction

Police use of lethal force has long been of interest to researchers. Early studies explored the incidence of and trends in police shootings, the demographic characteristics of the parties involved, and the events and circumstances leading up to each shooting (e.g., Fyfe, 1982; Geller and Karales, 1981; Milton, Halleck, Lardner, and Albrecht, 1977). More recently, in the wake of several high-profile shootings of unarmed minority suspects by white police officers, researchers have explored whether implicit (or even explicit) bias influences officer-involved shooting (OIS) decisions (e.g., James, James, and Vila, 2016; Klinger, Rosenfeld, Isom, and Deckard, 2016; Nix, Campbell, Byers, and Alpert, 2017; Tregle, Nix, and Alpert, 2019). This latest line of research has made important contributions, but many questions remain unanswered.

One important unanswered question is, what explains apparent differences between macro- and micro-level studies of police shooting decisions? For example, a county-level study by Ross (2015) found that, per unit population, unarmed black suspects were 3.5 times more likely to be shot by police than unarmed white suspects. In contrast, a subsequent incident-level study by Fryer (2019) found that, per unit

CONTACT John L. Worrall worrall@utdallas.edu

encounter, black civilians were less likely than whites to be shot. Indeed, several other incident- and officer-level studies have found that blacks are less likely to be shot relative to those in other racial/ethnic groups (e.g., Worrall, Bishopp, Zinser, Wheeler, and Phillips, 2018; Wheeler, Phillips, Worrall, and Bishopp, 2018). Macro-level studies (e.g., Klinger et al., 2016; Legewie and Fagan, 2016) tend to find no such relationships, or at least much weaker ones.

A nascent area of research has sought to unpack these seemingly contradictory, even paradoxical relationships. For example, Ross, Winterhalder, and McElreath (2018) developed a model to show that "reversed" racial disparities in recent work can be explained when accounting for "…consistent and systemic biases against black individuals" (p. 61) that occur well before force encounters. In other words, "…even if police do not show racial bias in the use of lethal force conditional on encounter, racial disparities in encounters themselves will still produce racial disparities in the population-level rates of the use of lethal force" (p. 69). More simply, there is heterogeneity in the rates at which various individuals come into contact with police. A problem is that researchers have not yet reached agreement on how best to accommodate this heterogeneity. Indeed, only recently have they even raised the issue in the OIS context (see, e.g., Cesario, Johnson, and Terrill, 2019; Johnson, Tress, Burkel, Taylor, and Cesario, 2019; Tregle et al., 2019).

The fundamental problem is benchmarking.[1] If we are to fully understand whether certain classes of persons are shot by police more often than others (or have force used against them more often), then we must know who is at risk of being shot and, importantly, *who is not*. The media are quick to point out, for example, that black citizens are disproportionately represented in OIS statistics relative to their numbers in the population (e.g., Swaine and McCarthy, 2017), yet it is well-known that not everyone has the same likelihood of coming into contact with the police. Prior research reveals that blacks are more likely than whites to be stopped (e.g., Smith, Rojek, Petrocelli, and Withrow, 2017) and arrested (Kochel, Wilson, and Mastrofski, 2011). As such, their risk of being involved in OIS is generally inflated. This has prompted researchers to search for better benchmarks to use in addressing the question of racial bias in officer-involved shootings.

Recent studies have used national OIS data to compare different benchmarks against one another (e.g., Cesario et al., 2019; Johnson et al., 2019; Tregle et al., 2019). Sample benchmarks used include population numbers, police-initiated contacts, traffic stops, pedestrian stops, arrests of various types, and even crime statistics. Each tells a different story. Benchmarking with arrests suggests much less racial disparity than benchmarking with Census data (e.g., Tregle et al., 209, p. 24). Similarly, when substituting population proportions with crime rates, anti-black disparities disappear (Cesario et al., 2019). These findings are also echoed in studies with benchmarks involving violence directed toward police (Shjarback and Nix, 2020). In contrast, traditional benchmarks drawn from Census data tend to suggest substantial racial disparities.

[1]Most scholarly attention to the issue has unfolded in the racial profiling literature (see, e.g., Alpert, Smith, and Dunham, 2004; Smith, Tillyer, Lloyd, and Petrocelli, 2019).

Recent OIS benchmarking work is interesting, timely, and helpful. It is mostly ignored, however, in incident-level OIS research—with two exceptions. The authors of two recent studies (Wheeler et al., 2018; Worrall et al., 2018) invoked a creative benchmarking approach by using shoot/don't shoot cases to explore bias in OIS incidents. Their approach excluded from the analyses interactions in which it was improbable someone would be shot given that officers never drew their weapons against them. The authors found that blacks were much less likely to be shot than whites, which contradicts the popular bias narrative perpetuated in the media—and also some recent findings in aggregate research, such as Ross (2015). As Tregle et al. (2019, p. 27) note, however, agency policies vary greatly from one to the next in terms of rules for officers drawing their weapons against suspects, making it unlikely this approach could work as a sort of uniform benchmark across a representative sample of police departments.

It is also possible that officers' decisions to draw their weapons may themselves be biased. This begs an important research question: Are police officers more or less likely to draw their weapons against minority suspects? The present study seeks to answer this question with 2017 use-of-force (UoF) data from the Dallas Police Department (DPD). DPD policy requires officers to report when they draw their weapons as part of UoF documentation procedures. It is thus possible to model police weapon draws in the same way shoot/don't shoot was modeled in the work of Wheeler et al. (2018) and Worrall et al. (2018). We begin with a model that does just this, then we estimate a series of secondary models limited to arrest cases and cases in which suspects acted aggressively[2] in an effort to limit interactions to those during which there would be an enhanced probability of suspects having weapons drawn against them.

Prior research

Deadly force and the question of bias

Decades of research reports on deadly force, also known as officer-involved shootings (OIS). As far back as 1963, Robin (1963) examined 32 police shootings in Philadelphia and found that victims were, on average, 27 years old and disproportionately African American. Additionally, most of the shootings occurred at night and outside, and most involved resistant suspects. In a later Chicago-based study, Geller and Karales (1981) reported that over 50 percent of shootings occurred because the suspect used or threatened to use a weapon (see also Milton et al., 1977). The authors also found that 80 percent of those shot were minorities, which was echoed in Fyfe's (1982) study. In contrast, though, Fyfe (1982) found that racial discrepancies were most prominent in "elective shootings of nonassaultive, unarmed people" (p. 721).

Most recently, research has incorporated the concept of implicit bias into OIS studies. Defined as "discriminatory biases based on implicit attitudes or implicit stereotypes" (Greenwald and Krieger, 2006, p. 951), implicit biases are subconscious aspects of perception, interpretation, and judgment that occur during the decision-making process. Applying this to police shootings, implicit bias manifests when officers

[2]There were advantages and disadvantages to each approach. We discuss them in the Results section.

subconsciously let preconceived ideas about certain individuals influence their decisions to use deadly force. And to the extent implicit bias exists in this context, subconscious influences can happen very quickly and imperceptibly, as shooting decisions are sometime made in terms of milliseconds (James, Vila, and Klinger, 2014; c.f. Fyfe, 1997). Implicit bias is thus problematic for two reasons. First, officers do not know it is occurring, and second, they have no time to counter it.

Whether implicit bias actually factors into shooting decisions, race is a central issue in officer-involved shootings, particularly in media accounts.[3] A number of minority communities also believe that bias dominates all aspects of criminal justice, not just shootings (e.g., Weitzer and Tuch, 2005). Certain data support the idea that minorities are shot much more often than whites relative to their representation in the population at large. In response to these concerns, researchers have recently sought to determine whether race is an objective determinate of OIS. Worrall et al. (2018) organized this research into five categories: "(1) aggregate-level research concerned with explaining the incidence of police shootings in specific geographic areas; (2) comparisons of cases involving different offense types; (3) shooting studies with non-shooting outcomes, such as whether the suspect was armed; (4) simulation studies; and (5) shoot/ don't shoot studies involving real-world samples" (p. 1176). We use their classification scheme to summarize recent research in this area.

Aggregate research

Aggregate-, or macro-level OIS research has itself fallen into three categories. At the national level, a flurry of recent studies has sought to determine whether blacks are shot disproportionately relative to whites. Cesario et al. (2019) benchmarked police shootings with a variety of crime statistics and found " … in nearly every case, [w]hites were either more likely to be fatally shot by police or police showed no significant disparity in either direction" (p. 588). When fatal shootings were benchmarked against crime rate estimates, they found "no systematic evidence of anti-Black disparities in fatal shootings, fatal shootings of unarmed citizens, or fatal shootings involving misidentification of harmless objects" (p. 586). They concluded that the anti-Black disparities seen at the per capita level are likely due to different rates of exposure to the police (as indexed by crime rates). Similar findings emerged from Tregle et al. (2019); when benchmarked against population proportions, the odds of black citizens being shot by police were nearly three times that for white, but when arrests, particularly arrests for weapons offenses, served as the benchmark, blacks were around two-thirds as likely as whites to be shot (see also Shjarback and Nix, 2020).[4]

---

[3]The Reverend Jessie Jackson called the Baton Rouge shooting of Alton Sterling "a legal lynching." (http://www.huffingtonpost.com/entry/alton-sterling-jesse-jackson-legal-lynching_us_577d126be4b0a629c1ab56c6). Another story in the *Huffington Post* characterized one OIS as an "extrajudicial" execution (http://www.huffingtonpost.com/justin_cohen/advice-for-white-folks-in_b_10861488.html).

[4]A reviewer raised the concern that Cesario et al. (2019) and Tregle et al. (2019) fail to address the shooting of unarmed people by the police, as there is no reason to assume they are violent criminals." In the reviewer's words, "[t]he excessive killing of unarmed non-aggressing black people is not justified on the basis of the average crime rate." We do not disagree with this point entirely, but it is fair to say "excessive" is open to interpretation because while, according to the work of Cesario et al. (2019), 40 unarmed blacks were killed in 2015 and 2016, 62 unarmed whites were also killed during the same period. The crime statistics used in Cesario et al. (2019) and Tregle et al. (2019) are *proxies* for enhanced exposure and they are obviously imperfect. The Shjarback and Nix (2020) study is something of an improvement given its benchmarking with violence against officers.

Another line of aggregate research looks at counties. Since there is variability in officer-involved shootings at this level, it is possible to model that variability with demographic data. In one of the first studies of this kind, Ross (2015) used county-level data to predict the relative risk of being shot by the police. He found evidence of significant bias in the killing of unarmed black Americans. Specifically, "the probability of being black, unarmed, and shot by the police is about 3.49 times the probability of being white, unarmed, and shot by the police" (p. 6).

Still other researchers have drilled down to the Census block level. For example, Klinger et al. (2016) modeled St. Louis police shootings using Census block group data. Key variables in their study included firearm violence, the percent black, median household income, percent owner-occupied houses, percent college graduates, and a spatial lag. Firearm violence was the most significant predictor of IOS. As for race, it " … does matter but only insofar as it increases the level of firearm violence and, even then, only to a point" (p. 212).

In a working paper, Legewie and Fagan (2016) emulated the work of Klinger et al. with Census designated places. They regressed 3,833 incidents of officer-involved shootings on various measures of group threat (proportion black, proportion black on city council, the white unemployment rate, and the black-on-white homicide rate) and minority representation in police departments. Results revealed that "black-on-white homicide rate is a significant predictor of officer-involved killings" (p. 32). They also found, however, that minority representation in the police department reduces OIS potential.

Case comparisons

Fryer (2019) studied officer-involved shootings from Houston between 2000 and 2015, yielding a total of 1,332 incidents. In one of his analyses, shooting cases were compared to serious arrest cases (which presumably could have ended in shootings). In another set of analyses, shooting cases were compared to TASER cases. The outcomes were dichotomized for analysis (shoot vs. arrest; shoot vs. TASER). Various suspect, officer, and encounter characteristics were then used in logistic regressions. Fryer found " … no racial differences in either the raw data or when contextual factors are taken into account" (p. 1210).[5] As of this writing, his work is the only study to invoke case comparisons.

Non-shooting outcomes

In a creative study, Nix et al. (2017) regressed two non-shooting variables on suspect race and other factors: (1) whether the suspect was not attacking the officer(s) or other civilians just before being shot and (2) whether the civilian was unarmed when fatally shot. In the first set of analyses, Nix et al. found that "other race" was significant, but "black" was not: "[C]ivilians from other racial/ethnic groups were significantly *more* likely than whites to have been in the non-attack group" (pp. 324-325). In the second set of analyses, they found that "black civilians were significantly *more* likely than white civilians to have been unarmed when they were shot and killed by the police" (p. 325). "Other race" was not significant in those analyses.

[5]A reviewer correctly pointed out that if there is racial bias not just in killing unarmed black people but also "tasing" black people, then the denominator would be inflated, giving a false impression of anti-white bias.

More recently, Johnson et al. (2019) avoided the benchmarking issue by attempting to predict the race of the person shot (see also Streeter, 2019). With civilian race as the outcome, they conducted a multinomial regression with various officer, civilian, and county characteristics as predictors. When county-level predictors were included in their models, "[b]lack officers were not more likely to fatally shoot [b]lack civilians" (p. 15879). Their approach, while interesting, changes the research question though and cannot reveal whether there is actual bias in *decisions to shoot*. Knox and Mummolo (2020) offer a critique of this work, and Knox, Lowe, and Mummolo (2019) extend a similar critique into the police investigation context. Nevertheless, it is an interesting contribution that clearly adds to the discussion in this area.

Simulation studies

A number of researchers have studied racial bias in laboratory settings. Early studies found that both police and non-police subjects are quicker to press a "shoot" button for black suspects relative to white suspects (e.g., Correll and Keesee, 2009; Correll, Park, Judd, and Wittenbrink, 2002, 2007; Correll, Park, Judd, Wittenbrink, Sadler, and Keesee, 2007; Sadler, Correll, Park, and Judd, 2012). There is evidence, however, that such tendencies decline over repeated trials (e.g., Plant and Peruche, 2005; Plant, Peruche, and Butz, 2005).

More recently, researchers have used state-of-the-art police training simulators to further explore the bias issue (James, Vila, and Daratha, 2013; James, Vila, and Klinger, 2014). This advance is timely, as Bennell and Jones (2005) argued that simulation training is capable of teaching officers' skills they may not be able to practice. Similarly, researchers have found that simulation training causes various physiological responses (e.g., elevated pulse and blood pressure) even though there is no real threat to the participants (e.g., Brisinda, Venuti, Cataldi, Efremov, Intorno, and Fenici, 2015; Brisinda, Fioravanti, Sorbo, Venuti, and Fenici, 2015; Johnson, Stone, Miranda, Vila, James, James, Rubio, and Berka, 2014; Winser, Hinson, James, Vila, Whitney, and Van Dongen, 2014).

Advanced simulation work by James et al. (2013, 2014) has found that police and non-police participants are slower to shoot armed black suspects than armed white suspects. More recently, James et al. (2016) exposed 80 officers to a total of 1,517 different scenarios. The outcomes they were most interested in were reaction time and whether an unarmed suspect was shot. Suspect race was the key predictor. Interestingly, officers took longer to shoot armed black suspects than armed white suspects. Their work was not immune from criticism. For example, Roussell et al. (2019) asserted that "… accepting such findings as indicators of lived experience may have dangerous consequences for police officers and the communities they police" (pp. 6-7).

Real-world shoot/don't shoot studies

Researchers have more recently examined actual shoot/don't shoot outcomes in a variety of contexts. Using data from the NYPD, Ridgeway (2016) paired shooting and non-shooting officers from the same scenes. Predictors included rank, years in the department, recruitment age, race, gender, education, and various performance indicators. He found that "black officers had more than three times greater odds of shooting than white officers" (p. 5). A key limitation with Ridgeway's study is that all the incidents ended in shootings. In other words, he did not compare shooting cases to cases in which the suspect *was not* shot.

In response to this and other limitations in OIS research, Worrall et al. (2018) compared OIS cases with similar cases in which a weapon was drawn but not fired. The study was conducted at the officer-level. Wheeler et al. (2018) conducted a similar study at the incident level, and also incorporated Census block demographic controls to address some of the concerns expressed by Ross et al. (2018). Both studies found that black suspects were much less likely to be shot than those in other racial/ethnic categories. In Worrall et al. (2018), blacks were about one-third as likely to be shot as suspects of other races and ethnicities. Wheeler et al. (2018) found that black suspects were approximately half as likely to be shot as white suspects (p. 61, Table 5). Both studies challenged the dominant implicit bias narrative, but it is possible at least one of the studies fell victim to concerns raised by Ross et al. (2018).

The novelty of the shoot/don't shoot approach notwithstanding, Wheeler et al. (2018) and Worrall et al. (2018) also adopted a unique benchmarking strategy. Indeed, their studies were perhaps the first to attempt benchmarking at the incident level. As mentioned earlier, the approach excluded cases in which it was unlikely someone would be shot because officers did not draw their weapons against them. Nevertheless, the approach was imperfect because it did not discern whether there is racial bias in officers' decisions to draw their weapons in the first place.

Current study

Despite the attention they receive, officer-involved shootings are still exceptionally rare. And as Ross et al. (2018) explain, this rarity makes statistical estimates particularly susceptible to error. Far more common are instances in which officers draw their weapons but do not fire. For example, Worrall et al. (2018) analyzed data from a police department that had just over 100 officer-involved shootings in a six-year period compared to 1,700 "weapon draws" in a shorter four-year period. Adding to this, prior research reveals that most officers do not use deadly force even when it is apparently justified (Pinizzotto, Davis, Bohrer, and Infanti, 2012). The study of non-shooting incidents, then, could help overcome problems in previous research, or at least improve the sampling of cases with which to examine the issue of implicit and explicit bias in police force decisions.

Our study seeks to add to past OIS and use-of-force research by modeling officers' decisions to draw their weapons as a function of suspect race and other suspect, officer, and incident characteristics. This approach is novel insofar as officers' decisions to shoot could be a function of their (potentially biased) decisions to draw their weapons in the first instance. Officer-involved shootings, and by extension applications of less-lethal force, cannot be viewed in isolation. It is necessary to move backward in time and contemplate the contingent and evolving nature of police use of force decision-making (Terrill, 2001).

Methods

We collected all use-of-force (UoF) data from the Dallas Police Department during 2017, yielding a sample size of 2,150 incidents. DPD policy states, in part, "Officers

Table 1. Summary statistics.

	Dichotomous Variables		Continuous Variables			
	1	0	Mean	Std. Dev.	Minimum	Maximum
Dependent Variable						
Weapon Drawn	389 (18%)	1,761 (82%)	–	–	–	–
Independent Variables						
Black Suspect	1,122 (52%)	1,028 (48%)	–	–	–	–
Black Officer	268 (12%)	1,882 (88%)	–	–	–	–
Male Officer	1,905 (89%)	245 (11%)	–	–	–	–
Male Suspect	1,795 (83%)	355 (17%)	–	–	–	–
Officer Tenure	–	–	7.46	6.87	0	36
Suspect Mentally Unstable	447 (21%)	1,703 (79%)	–	–	–	–
Suspect UI Drugs	1,133 (53%)	1,017 (47%)	–	–	–	–
Risky Call	199 (9%)	1,951 (91%)	–	–	–	–
Suspect Active Aggression	370 (17%)	1,780 (83%)	–	–	–	–
Suspect Weapon Display	149 (7%)	2,001 (93%)	–	–	–	–

Notes: Risky Call = call for cover, crime in progress, suspicious activity, and/or warrant execution. Observations = 2,150 for gun draw models; 1,761 for TASER draw models.

may draw or display firearms when there is a threat or reasonable belief that there is a threat to life or they have a reasonable fear for their own safety and/or the safety of others."[6] For weapon draw cases, we limited our analyses strictly to those instances in which officers drew their service weapons.[7]

As reported in Table 1 below, there were 389 weapon draws during 2017 compared to 1,761 other force incidents. Other types of force included joint locks, physical restraint, TASER display/deployment, take-downs, and strikes with hands, feet, or impact weapon (e.g., ASP baton).

Variables and coding

The main dependent variable in our analyses was weapon drawn/other force. "Weapon drawn" cases were coded with a 1 and "other force" cases were coded with a 0. The key predictor in our analyses was suspect race/ethnicity (black = 1; other = 0).[8] Suspect gender was coded with a 1 for male, 0 for female. Officer characteristics included race (black = 1; other = 0), gender (male = 1; female = 0), and years of experience at the time of the incident ("officer tenure," a count).

Various combinations of five incident characteristics were also included in the analyses. Two of them, "suspect mentally unstable" and "suspect under the influence of drugs," were evaluated by the officer(s) on the scene at the time of the force incident. A "risky call" variable denoted incidents in which force was perhaps more likely. Risky calls included the following: call for cover, crime in progress, suspicious activity, and/or warrant execution. Next, "suspect active aggression" was defined, per DPD policy 901.00 as "physical actions of assault toward officer. This includes a suspect who attempts to attack, grabs, punches, kicks, and/or wrestles with the officer, or a suspect

[6]Dallas Police Department General Order 906.02 (np). Rev. 6/19/09.
[7]Dallas Police Department training, and thus required reporting, does not distinguish between drawing and pointing. No officers are taught, for instance, to draw their weapons and hold them at their sides without pointing them at someone or something.
[8]We adopted different coding schemes in supplementary analyses reported below.

Table 2. Logistic regression results (Gun draw outcome).

Variable	Suspect Characteristics	Officer Characteristics	Full Model
Black Suspect	0.68** (0.52, 0.90)	– –	0.67** (0.51, 0.89)
Black Officer	– –	0.73 (0.51, 1.06)	0.67 (0.42, 1.06)
Male Officer	– –	0.85 (0.61, 1.19)	1.26 (0.77, 2.04)
Male Suspect	5.76*** (3.06, 10.83)	– –	5.30*** (2.81, 9.97)
Officer Tenure	– –	0.98* (0.96, 1.00)	0.97* (0.95, 1.00)
Suspect Mentally Unstable	0.21*** (0.13, 0.31)	– –	0.21*** (0.14, 0.32)
Suspect UI Drugs	0.23*** (0.17, 0.31)	– –	0.22*** (0.17, 0.30)
Risky Call	– –	– –	1.27 (0.83, 1.95)
Suspect Active Aggression	0.42*** (0.26, 0.68)	– –	0.44** (0.27, 0.70)
Suspect Weapon Display	49.34*** (26.96, 90.31)	– –	54.46*** (29.25, 101.75)
Constant	0.11 (0.05, 0.21)	0.31 (0.22, 0.43)	0.12 (0.05, 0.26)
Nagelkerke R-Square	0.403	.01	0.409

Notes: Odds ratios reported (95% confidence intervals in parentheses). $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Observations = 2,150 (i.e., full sample).

that displays an obvious imminent intent to attack, grab, punch, kick, and/or wrestle with the officer." Finally, "suspect weapon display" refers to a suspect presenting any type of weapon, or perceived to be armed with a weapon, *and* being perceived as an immediate threat. A weapon does not necessarily need to be a gun.[9]

Modeling strategy

We first provide descriptive analysis of the differences between each of the variables in the weapon "draw" and "no draw" control cases. Next, we fit two logistic regression models predicting the probability of a weapon being drawn on combinations of different predictors. The full model was:

$$\text{Prob}(\text{Draw}) = f\left[\beta(\text{Suspect}) + \gamma(\text{Officer}) + \psi(\text{Incident})\right] \quad (1)$$

In this model, *f* is the logistic function. For example, one can assess the probability that a black suspect is more likely to have a weapon drawn against them than someone of a different race/ethnicity, adjusting for whether the suspect displayed a weapon or acted in a threatening manner toward officers.

Results

Table 2 presents the odds ratios for three separate logistic regression models with the 1,0 gun draw variable as the outcome. The first model consists solely of suspect characteristics. Black suspects were approximately two-thirds as likely as non-blacks to have a weapon drawn against them. Other significant variables included male suspect, suspect mentally unstable, suspect under the influence of drugs, suspect actively aggressive, and suspect weapon display. The latter two findings are interesting. Actively aggressive suspects were less than half as likely as non-aggressive suspects to have weapons drawn against them. One plausible explanation is that officers in such situations were not given an opportunity to draw their weapons, perhaps because the suspect was in the midst of attacking, grabbing, punching, kicking, and/or wrestling

[9]It would be ideal to have a deadly resistance category, but no such indicator is available in the official data.

with the officer. Alternatively, and perhaps more likely, is that while such resistance by definition is aggressive, officers respond with less-lethal proportional force.[10] As for suspect weapon displays, officers were nearly 50 times as likely to draw their weapons against suspects who displayed a weapon.

The second column of odds ratios in Table 2 consists of officer characteristics, with only one significant effect—officers with more experience on the job were less likely to draw their weapon. The full model in the third column combines both suspect and officer characteristics. Results were consistent with the suspect-specific findings reported in the first column. Once again, the key takeaway from Table 2 is that black suspects were less likely to have weapons drawn against them relative to those in other racial/ethnic groups. This finding could be an artifact of no benchmarking, which we address below.

In order to put the gun draw findings in context, we conducted another set of analyses in which a TASER draw served as the outcome. All gun draw cases were excluded from such models, as it was rare for officers to draw guns and TASERs in the same incidents.[11] Recall from Table 1 that there were 1,761 cases of other types of force being used during the observation period. Of those, 179 were TASER draws, with the remaining 1,582 cases consisting of other types of force short of drawing a weapon. Thus, the number of cases used to estimate the models in Table 3 was 1,761. Note two interesting findings in Table 3. First, the black suspect variable failed to reach statistical significance. Black suspects were no more or less likely than others to have TASERs drawn against them. Second, fewer predictors were significant in Table 3 relative to Table 2, including active resistance, suggesting that officers are motivated to draw their TASERs for different reasons than their weapons.

The benchmarking problem

As noted earlier, a critical issue in assessing police use of lethal force is benchmarking (Cesario et al., 2019; Tregle et al., 2019). A commonly used benchmark has been the use of population ratios (Brown and Langan, 2001; Gabrielson, Sagara, and Jones, 2014; Takagi, 1974), in which researchers compare how often suspects of different races are fatally shot by the police given their population proportions. The problem is this approach presumes that the opportunity for police lethal force is equally likely for every person within each group. However, prior research has shown that the use of lethal force is strongly tied to crime-related contexts, with the modal police shooting being one in which suspects pose a potentially deadly threat (Binder and Fridell, 1984;

[10]Yet another plausible explanation for this finding is that active aggression and suspect weapon display are collinear, or that one is a subset of other. Fortunately, this was not a possibility in our data. Officers could choose either active aggression or suspect weapon display as their reason for using force, but not both. Nevertheless, we estimated separate models excluding one or the other while keeping all the other variables listed in the last column on table 2. When active aggression was retained and weapon display was excluded, the active aggression odds ratio was .29 ($p < .001$), and when weapon display was retained and active aggression was excluded, weapon display odds ratio was 60.89 ($p < .001$). No other findings were appreciably altered.

[11]If multiple force "types" are deployed in a particular encounter, which is common, officers are required to report what was used first, second, and so on. In the 2017 UoF data, there were just two cases in which a TASER was followed by a gun draw, and there were just 15 cases where a weapon was drawn first, then followed by a TASER. In our TASER models, we excluded *all* cases in which a gun was drawn, including the 17 just mentioned.

Table 3. Logistic regression results (TASER draw outcome).

Variable	Suspect Characteristics	Officer Characteristics	Full Model
Black Suspect	1.08 (0.78, 1.49)	– –	1.09 (0.79, 1.51)
Black Officer	– –	0.97 (0.60, 1.55)	0.81 (0.49, 1.35)
Male Officer	– –	1.03 (0.63, 1.69)	1.05 (0.62, 1.77)
Male Suspect	1.71* (1.07, 2.73)	– –	1.69* (1.06, 2.70)
Officer Tenure	– –	0.97* (0.95, 1.00)	0.97* (0.94, 0.99)
Suspect Mentally Unstable	0.84 (0.53, 1.32)	– –	0.86 (0.54, 1.36)
Suspect UI Drugs	0.65* (0.44, 0.96)	– –	0.63** (0.42, 0.93)
Risky Call	– –	– –	0.71 (0.38, 1.33)
Suspect Active Aggression	1.13 (0.76, 1.67)	– –	1.16 (0.78, 1.71)
Suspect Weapon Display	21.55*** (7.12, 65.20)	– –	26.28*** (8.17, 84.53)
Constant	0.08 (0.05, 0.15)	0.14 (0.08, 0.22)	0.11 (0.05, 0.23)
Nagelkerke R-Square	0.04	0.005	0.04

Notes: Odds ratios reported (95% confidence intervals in parentheses). $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Gun draw cases excluded. Observations = 1,760.

Binder and Scharf, 1980; Geller and Karales, 1981; Koper, 2016; Fyfe, 1980, 1981; Selby, Singleton, and Flosi, 2016; White, 2016). Insofar as suspects of varying races have different police exposure rates, a more appropriate benchmark is the rate of police exposure, which may vary across groups (Barnes, Jorgensen, Beaver, Boutwell, and Wright 2015).

Three recent national-level studies (Cesario et al., 2019; Shjarback and Nix, 2020; Tregle et al., 2019), discussed earlier, have sought to address this issue using benchmarks seeking to account for such differential exposure. Cesario et al. (2019) found blacks were more likely than whites to be killed by the police when using population as a benchmark, but not when benchmarked against reported criminal activity. Shjarback and Nix (2020) used killings of and assaults against officers as benchmarks and found that blacks were fatally shot almost half as often as whites. Tregle et al. (2019) found blacks were less likely to be killed by the police when benchmarked against weapon and violent crime arrests.

Arrest benchmark

Given the importance of benchmarking, we estimated a third set of models limiting observations to cases in which suspects were arrested. Whereas Wheeler et al. (2018) and Worrall et al. (2018) excluded interactions in which it was improbable someone would be shot given that officers never drew their firearms, by limiting our analyses to cases in which suspects were arrested, we excluded interactions in which it is less likely suspects would have had weapons drawn against them given that officers never arrested them. A two-way crosstabulation of arrests versus gun draws appears in Table 4 (we also include separate crosstabs broken down by suspect race).[12]

The arrest benchmark is defensible for both practical and empirical reasons. In practical terms, arrest cases offer a valid benchmark because they consist of interactions with police that may have more potential than non-arrests to turn deadly (Cesario et al., 2019). Moreover, the potential for resistance, which in turn could lead to an officer to draw a weapon, or even possibly fire it, is higher in arrest relative to non-arrest encounters. In an early study, Crawford and Burns (2002) found that 25 percent of

[12]Thirty percent of the non-arrest cases involved weapon draws, so it is fair to say this benchmarking approach is imperfect. As such, we conducted additional analyses (see below) limited to active aggression cases.

Table 4. Arrest vs. Gun Draw cross tabulation.

All Cases

	No weapon	Weapon drawn
No Arrest	138	60
Arrest	1,623	329

Black Suspects

	No weapon	Weapon drawn
No Arrest	80	29
Arrest	848	165

Other Suspects

	No weapon	Weapon drawn
No Arrest	58	31
Arrest	775	164

arrestees in their Phoenix sample resisted in some form or fashion. In our data from Dallas, 27 percent of arrestees resisted; no suspects who were not arrested resisted.

Benchmarking in this context is concerned fundamentally with who is at risk for a particular action. In our study, we must know which suspects are at an elevated risk of having a weapon drawn against them but do *not* have a weapon drawn against them. Limiting a subset of analyses to arrest cases helps address this problem because not all individuals against whom the police report using force are at the same risk of having a weapon drawn against them. The odds ratio approach used in other studies (e.g., Tregle et al., 2019; Cesario et al., 2019) is not helpful (for us) because we are not concerned solely with the issue of race, but rather with whether and how race is associated with weapon draws while controlling for other factors that may affect weapon draws. For example, we could have calculated an odds ratio to explore racial bias in gun draws vis-à-vis arrests, but we would not have been able to control for other important incident-level covariates. Finally, controlling for arrest in models such as those reported in Table 2 does not solve the benchmarking problem because the sample would still contain people who were unlikely to have weapons drawn against them (i.e., those who were not arrested).

Results from additional logistic regressions, limited to arrest cases, appear in Table 5. Several interesting findings emerged. First, and most noteworthy, the black suspect variable failed to reach statistical significance. That is, once the arrest benchmark was incorporated, black suspects were no more or less likely than others to have weapons drawn against them. This effect was observed across the suspect characteristics, officer characteristic, and full models. The effect was, however, significant at the .10 level (not shown in Table 5).

This finding is difficult to fully explain when compared to the full sample of all force cases (i.e., results reported in Table 2). In one sense, by restricting cases to only those involving arrests one may posit that race would not have an effect because of the enhanced threat likelihood, which was what was found as reported in Table 5. That is, officers view the enhanced potential threat more readily, which overrides racial distinctions. Yet, a racial distinction was found when considering all force cases, with black suspects being *less* likely to be subject to a weapon draw. Such a finding

Table 5. Logistic regression results (Gun draw outcome with arrest benchmark).

Variable	Suspect Characteristics	Officer Characteristics	Full Model
Black Suspect	0.77 (0.57, 1.03)	– –	0.75 (0.56, 1.01)
Black Officer	– –	0.75 (0.15)	0.71 (0.44, 1.16)
Male Officer	– –	0.88 (0.16)	1.39 (0.81, 2.40)
Male Suspect	6.01*** (3.01, 12.01)	– –	5.37*** (2.69, 10.72)
Officer Tenure	– –	0.96*** (0.01)	0.96** (0.94, 0.99)
Suspect Mentally Unstable	0.26*** (0.17, 0.39)	– –	0.26*** (0.17, 0.41)
Suspect UI Drugs	0.26*** (0.19, 0.36)	– –	0.25*** (0.18, 0.35)
Risky Call	– –	– –	1.20 (0.76, 1.91)
Suspect Active Aggression	0.40*** (0.24, 0.66)	– –	0.41** (0.24, 0.69)
Suspect Weapon Display	59.97*** (29.99, 119.92)	– –	68.75*** (33.23, 142.24)
Constant	0.09 (0.04, 0.18)	0.31 (0.06)	0.10 (0.04, 0.23)
Nagelkerke R-Square	0.38	0.02	0.39

Notes: Odds ratios reported (95% confidence intervals in parentheses). *$p < .05$; **$p < .01$; ***$p < .001$. Observations = 1,952.

suggests that officers viewed black suspects as less of a potential danger overall. One can only speculate on such findings. Perhaps officers in a post-Ferguson climate are more cautious when drawing their firearms during interactions with black suspects for fear of scrutiny. Alternatively, perhaps officers overestimate the potential threat white suspects pose and thus draw their weapons more readily. Either way, such a finding mirrors much of the national debate about race and police use of force. It is important to keep in mind that while the statistical significance varied across the two analyses, officers were less likely to draw their weapon in both.

Focusing on the full model, other interesting findings emerged. Concerning tenure, the longer officers served, the less likely they were to draw their weapons against black suspects in arrest cases, but the odds ratio was close to unity. Similar effects were observed in Table 5 relative to Table 2 for the following variables: male suspect, suspect mentally unstable, suspect under the influence of drugs, suspect active aggression, and suspect weapon display. Benchmarking did little to affect the odds ratios for these variables, and significance levels remained more or less consistent. All in all, the arrest benchmark did little to change the results *apart from the race issue*.

Alternative race coding

To ensure that the race results in Table 5 were not an artifact of our race coding scheme, we estimated some supplementary models using different coding schemes and sample restrictions. First, we restricted the sample to black and white suspects. With the arrest benchmark this reduced the sample size to 1,491 (without the arrest restriction, the full sample dropped from 2,150 to 1,643). Results did not differ from Table 5 (with the exception of officer tenure, which lost significance). Again, black suspects were no more or less likely than whites to have weapons drawn against them. Second, we dropped all but white, black, and Hispanic from the sample and created a "minority" variable coded as 1 for black or Hispanic and 0 for white. This reduced the number of cases in the model from 1,952 to 1,935 (from 2,150 to 2,133 in the full sample). As coded, minorities were no more or less likely to have weapons drawn against them than whites. The results from these models are reported in the first two columns of odds ratios in the Appendix.

Active aggression benchmark

Comparing gun draws to all other force types and/or limiting analyses to arrest cases (recall 30 percent of non-arrest cases involved weapon draws) could present problems. For example, a reviewer noted that the analyses reported thus far compared incidents that were serious enough to result in weapon draws to those potentially "benign" incidents in which it presumably was not. In other words, it is possible that some of the suspects in our preceding analyses had a zero probability of having weapons drawn against them. To address this problem, we estimated supplementary models in which we limited the observations to active aggression cases rather than arrest cases. Results appear in the last column of the table in the Appendix.

The results in the last column of the Appendix also help address another potential problem. Namely, if there is bias what causes arrest, our findings could mistakenly reveal anti-white bias (see, e.g., Ross et al., 2018). By restricting our analyses to active aggression cases, then, we overcome that problem to some extent. As shown in the Appendix, though, the sample dropped off considerably. We thus obtained arrest and reported offense data for several serious offenses and calculated odds ratios identical to those in Tregle et al. (2019, p. 23). We found that, in Dallas in 2017, blacks were just 1.25 times as likely as whites to be arrested for the serious offenses of assault, burglary, homicide, and robbery. Had we relied on Census representation in the denominator, the odds ratios would have been in excess of eight.

Discussion

A flurry of recent scholarship has examined the nature and extent of bias in officer-involved shootings. The main question asked is, are black suspects more likely than white suspects to be shot by police? Researchers diverge in their answers to this question. Incident-level studies (e.g., Wheeler et al., 2018; Worrall et al., 2018) tend to find that black suspects are less likely than whites and people in other racial ethnic/categories to be shot. Macro-level studies at the Census block (e.g., Klinger et al., 2016), county (e.g., Ross, 2015), and national (e.g., Tregle et al., 2019) levels tend to find either null or opposite effects.

Recent benchmarking research (e.g., Cesario et al., 2019; Johnson et al., 2019; Tregle et al., 2019) has sought to make sense of and unpack these seemingly contradictory relationships, but the work is largely national in scope. Only two studies (Wheeler et al., 2018; Worrall et al., 2018) appear to have adopted benchmarking procedures to address the question of racial bias at the incident level. Those studies excluded cases in which it was unlikely someone would be shot insofar as officers did not draw weapons against them. The problem with both studies, however, was that they did not address the question of whether officers are more inclined to draw their weapons against black suspects.

Our study sought to address this limitation while also invoking two distinct benchmarking procedures. We gathered 2017 Use-of-Force data from the Dallas police department and explored whether and to what extent various officer-, suspect-, and incident-level variables were associated with officers' decisions to draw their weapons. Dallas UoF reporting procedures require that officers document weapon draws, so it was possible—at least in Dallas—to back up in time from the point of a shooting to the initial weapon draw. For comparison, we modeled weapon draws and TASER draws. We also benchmarked our analyses with arrest cases. That is, in a third set of analyses we excluded interactions in which it was less likely someone would have a gun drawn against them given that officers never arrested them.

We found that, without any attempt to benchmark, blacks were approximately two-thirds as likely as others to have weapons drawn against them, net of other factors (see Table 2). These results held up regardless of the race coding scheme used. Similar effects were *not* observed in TASER models (Table 3). When we benchmarked by excluding cases in which an arrest was not made, the race effect disappeared. That is, blacks were no more or less likely than others to have guns drawn against them, at least at the .05 significance level. And this effect was observed after controlling for a number of other important potential correlates of weapon draws, including officer demographics, incident characteristics, and whether the suspects behaved aggressively and/or displayed weapons themselves.

A second approach to benchmarking limited our analyses to active aggression cases. This approach was taken, in part, because 30 percent of the non-arrest cases involved weapon draws. While limiting to active aggression cases restricted the sample considerably, the race variable remained insignificant. Specifically, blacks were no more or less likely to have weapons drawn against them.

Our key finding is somewhat unique relative to recent research in this area. On the one hand, recent national benchmarking studies tend to find that blacks are less likely to be shot than whites (e.g., Cesario et al., 2019; Tregle et al., 2019). The same was observed in some incident-level work (Wheeler et al., 2018; Worrall et al., 2018). On the other hand, some county-level work (Ross, 2015) has found that blacks are much more likely than whites to be shot. Our findings land somewhere in the middle. How do we reconcile these findings? The national level studies have not examined incident characteristics. The same holds with other aggregate research, whether at the county, Census block, or some other level. To our knowledge, our study is the first to attempt benchmarking at the incident level—with a focus on weapon draws. It will be interesting to see if others pick up the torch and explore the weapon draw issue further.

Limitations

We have pointed out various limitations throughout the study, but some key ones should be highlighted here once again. Most importantly, our work is limited to a single department. Policies governing weapon draws obviously vary across departments, so the outcome variable in Dallas may be qualitatively different than a similar outcome in another agency. Next, while benchmarking with arrests and active aggression cases improves over several recent incident-level studies, many of which fail to attempt any benchmarking at all, it is imperfect. For example, it fails to account for the possibility that officers are biased in their decisions to confront certain classes of suspects in the first place (to address this, we estimated additional models discussed in the Appendix—see above).

There is also no clear guidance on what constitutes the appropriate list of covariates to include in a model of police weapon draws. Our predictors were selected largely out of convenience because they were available in the data. Finally, the UoF and suspect data were all reported post-incident. For example, officers could have indicated a higher level of resistance to justify the force in question.

We should also take this opportunity to point out one of the most glaring weaknesses of this and *all other* OIS research: No study has yet explored the issue of racial bias in justified deadly force cases compared to those in which deadly force could have been used (i.e., was legally justified) but was not. For research in this area to move forward, it will be necessary for police departments to collect data on this important outcome.

Conclusion

Benchmarking is critical in any examination of racial bias in criminal justice decision making. So far this has not trickled down to OIS research as much as it should. This study included, just three papers have explored the issue with incident- and/or officer/suspect-level data. It appears that an arrest benchmark erases the appearance of racial bias, at least in a weapon drawing context. Obviously the arrest benchmark would not work in a strictly deadly OIS context because those who are shot and killed cannot be arrested. We thus encourage future researchers to look more closely at the question of weapon draws and use of force. Because there are so many more of these incidents than shootings, there are many opportunities to further probe the bias issue. It would be particularly instructive for researchers to consider other benchmarks besides aggregate arrests. For example, benchmarking at the incident level with specific arrest types (a la Tregle et al., 2019) would be constructive. Our data did not afford us the opportunity to do that.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

John L. Worrall is Professor of Criminology and Criminal Justice at the University of Texas at Dallas. He has published articles and book chapters on a variety of topics ranging from legal issues in policing to crime measurement. He is also the author or coauthor of 17 books, including the popular *Crime Control in America: What Works?* (4th ed., Pearson) and *Introduction to Criminal Justice* (16th ed., Cengage). He currently serves as Editor of the journal *Police Quarterly* and as Executive Director of the Academy of Criminal Justice Sciences.

Stephen A. Bishopp holds an appointment as Adjunct Assistant Professor (w/o Tenure)-University of Texas School of Public Health. He is also a Sergeant and 30-year veteran of the Dallas Police Department. He received his PhD (2013) in Criminology from the University of Texas at Dallas. His research interests include police mental health/well-being, police behavior, and criminological theory.

William Terrill is Interim Associate Dean in the Watts College of Public Service and Community Solutions, and Professor in the School of Criminology and Criminal Justice at Arizona State University. His research centers on police behavior, with an emphasis on police use of force and police culture. He has published numerous scholarly articles, chapters, and reports, as well as two books entitled "Police Coercion: Application of the Force Continuum" (2001, LFB Scholarly Publishing) and "Police Culture: Adapting to the Strains of the Job" (2013, Carolina Academic Press).

References

Alpert, G. P., Smith, M. R., & Dunham, R. G. (2004). Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research. *Justice Research and Policy*, *6*(1), 43–69. doi:10.3818/JRP.6.1.2004.43

Barnes, J. C., Jorgensen, C., Beaver, K. M., Boutwell, B. B., & Wright, J. P. (2015). Arrest prevalence in a national sample of adults: The role of sex and race/ethnicity. *American Journal of Criminal Justice*, *40*(3), 457–465. doi:10.1007/s12103-014-9273-3

Bennell, C., & Jones, N. J. (2005). *The effectiveness of use of force simulation training final report*. Ottowa, Canada: Psychology Department, Carleton University.

Binder, A., & Fridell, L. (1984). Lethal force as a police response. *Criminal Justice Abstracts*, *16*, 250–280.

Binder, A., & Scharf, P. (1980). The violent police-citizen encounter. *The Annals of the American Academy of Political and Social Science*, *452*(1), 111–121. doi:10.1177/000271628045200111

Brisinda, D., Venuti, A., Cataldi, C., Efremov, K., Intorno, E., & Fenici, R. (2015). Real-time imaging of stress-induced cardiac autonomic adaptation during realistic force-on-force police scenarios. *Journal of Police and Criminal Psychology*, *30*(2), 71–86. doi:10.1007/s11896-014-9142-5

Brisinda, D., Fioravanti, F., Sorbo, A. R., Venuti, A., & Fenici, R. (2015). Psychophysiological assessment of acute stress induced by high-pressure law-enforcement driving: A pilot study. *Psychology and Social Behavior Research*, *2*, 36–50.

Brown, J. M., & Langan, P. A. (2001). *Policing and homicide 1976–98: Justifiable homicide by police, police officers murdered by felons. Bureau of Justice Statistics*. Washington, DC: U.S. Department of Justice.

Cesario, J., Johnson, D. J., & Terrill, W. (2019). Is there evidence of racial disparity in police use of deadly force? Analyses of officer-involved fatal shootings in 2015-2016. *Social Psychological and Personality Science*, *10*(5), 586–595. doi:10.1177/1948550618775108

Correll, J., & Keesee, T. (2009). Racial bias in the decision to shoot?. *The Police Chief*, *76*, 54–57.

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*(6), 1314–1329.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007). Stereotypes and racial bias in the decision to shoot. *European Journal of Social Psychology*, *37*(6), 1102–1117. doi:10.1002/ejsp.450

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). 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. doi:10.1037/0022-3514.92.6.1006

Crawford, C., & Burns, R. (2002). Resisting arrest: Predictors of suspect non-compliance and use of force against police officers. *Police Practice and Research*, *3*(2), 105–117. doi:10.1080/15614260290033611

Fyfe, J. J. (1980). *Philadelphia police shootings 1975–78: A system model analysis. Civil rights division*. Washington, DC: U.S. Department of Justice.

Fyfe, J. J. (1981). Toward a typology of police shootings. In James J. Fyfe (Ed.), *Contemporary issues in law enforcement* (pp. 136–151). Beverly Hills, CA: Sage.

Fyfe, J. J. (1982). Blind justice: Police shootings in Memphis. *The Journal of Criminal Law and Criminology (1973-)*, *73*(2), 707–722. doi:10.2307/1143112

Fyfe, J. J. (1997). The split-second syndrome and other determinants of police violence. In Roger G. Dunham & Geoffrey Alpert (Eds.), *Critical Issues in Policing: Contemporary Readings* (3rd ed., 531–546). Prospect Heights, IL: Waveland Press.

Fryer, R. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, *127*(3), 1210–1261. doi:10.1086/701423

Gabrielson, R., Sagara, E., & Jones, R.G. (2014). Deadly force, in black and white: A pro publica analysis of killings by police shows outsize risk for young black males. *Pro Publica*. Retrieved from https://www.propublica.org/article/deadly-force-in-black-and-white.

Geller, W. A., & Karales, K. J. (1981). Shootings of and by Chicago police: Uncommon crises Part I: Shootings by Chicago police. *The Journal of Criminal Law and Criminology (1973-)*, *72*(4), 1813–1866. doi:10.2307/1143256

Greenwald, A. G., & Krieger, L. H. (2006). "Implicit bias: Scientific foundations. *California Law Review*, *94*(4), 945–967. doi:10.2307/20439056

James, L., James, S. M., & Vila, B. J. (2016). The reverse racism effect. *Criminology & Public Policy*, *15*(2), 457–479. doi:10.1111/1745-9133.12187

James, L., Vila, B. J., & Daratha, K. (2013). The influence of suspect race and ethnicity on decisions to shoot in a deadly force judgment and decision-making simulator. *Journal of Experimental Criminology*, *9*(2), 189–212.

James, L., Vila, B. J., & Klinger, D. (2014). Racial and ethnic bias in decision to shoot seen through a stronger lens: Experimental results from high-fidelity laboratory simulations. *Experimental Criminology*, *10*, 232–240.

Johnson, D. J., Tress, T., Burkel, N., Taylor, C., & Cesario, J. (2019). Officer characteristics and racial disparities in fatal officer-involved shootings. *Proceedings of the National Academy of Sciences*, *116*(32), 15877–15882. doi:10.1073/pnas.1903856116

Johnson, R. R., Stone, B. T., Miranda, C. M., Vila, B. J., James, L., James, S. M., Rubio, R. F., & Berka, C. (2014). Identifying psychophysiological indices of expert vs. novice performance in deadly force judgment and decision making. *Frontiers in Human Neuroscience*, *8* Epub ahead of print: doi:10.3389/fnhum.2014.00512. https://psycnet.apa.org/record/2014-49982-001

Klinger, D., Rosenfeld, R., Isom, D., & Deckard, M. (2016). Race, crime, and the micro-ecology of deadly force. *Criminology & Public Policy*, *15*(1), 193–222. doi:10.1111/1745-9133.12174

Knox, D., & Mummolo, J. (2020). Making inferences about racial disparities in police violence. *Proceedings of the National Academy of Sciences*, *117*(3), 1261–1262.

Knox, D., Lowe, W., & Mummolo, J. (2019). Administrative records mask racially biased policing. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3336338.

Kochel, T. R., Wilson, D. B., & Mastrofski, S. D. (2011). Effect of suspect race on officers' arrest decisions. *Criminology*, *49*(2), 473–512. doi:10.1111/j.1745-9125.2011.00230.x

Koper, C. S. (2016). Advancing research and accountability on police use of deadly force. *Criminology & Public Policy*, *15*(1), 187–191. doi:10.1111/1745-9133.12192

Legewie, J., & Fagan, J. (2016). *Group threat, police officer diversity and the deadly use of police force* (Columbia Public Law Research Paper No. 14-512). Retrieved from ssrn.com/abstract=2778692.

Milton, C. H., Halleck, J. W., Lardner, L., (1977). and G. L. & Abrecht, *Police use of deadly force*. Washington, DC: Police Foundation.

Nix, J., Campbell, B. A., Byers, E. H., & Alpert, G. P. (2017). A bird's eye view of civilians killed by police in 2015: Further evidence of shooter bias. *Criminology and Public Policy*, *16*, 1–32.

Pinizzotto, A. J., Davis, E. F., Bohrer, S. B., & Infanti, B. J. (2012). Law enforcement restraint in the use of deadly force within the context of 'the deadly mix'. *International Journal of Police Science & Management*, *14*(4), 285–298. doi:10.1350/ijps.2012.14.4.289

Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, *16*(3), 180–183. doi:10.1111/j.0956-7976.2005.00800.x

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*(2), 141–156. doi:10.1016/j.jesp.2004.07.004

Ridgeway, G. (2016). Officer risk factors associated with police shootings: A matched case-control study. *Statistics and Public Policy*, *3*(1), 1–6. doi:10.1080/2330443X.2015.1129918

Robin, G. D. (1963). Justifiable homicide by police officers. *The Journal of Criminal Law, Criminology, and Police Science*, *54*(2), 225–231. doi:10.2307/1141171

Ross, C. T., Winterhalder, B., & McElreath, R. (2018). Resolution of apparent paradoxes in the race-specific frequence of us-of-force by police. *Palgrave Communications*, *4*(1), 61–70. doi:10.1057/s41599-018-0110-z

Ross, C. T. (2015). A multi-level Bayesian analysis of racial bias in police shootings at the county-level in the United States, 2011–2014. *PloS One*, *10*(11), e0141854. doi:10.1371/journal.pone.0141854

Roussell, A., Henne, K., Glover, K. S., & Willits, D. (2019). Impossibility of a "reverse racism" effect: A rejoinder to James, James, and Vila. *Criminology & Public Policy*, *18*(1), E5–16. doi:10.1111/1745-9133.12289

Sadler, M. S., Correll, J., Park, B., & Judd, C. M. (2012). The world is not black and white: Racial bias in the decision to shoot in a multiethnic context. *Journal of Social Issues*, *68*(2), 286–313. doi:10.1111/j.1540-4560.2012.01749.x

Selby, N., Singleton, B., & Flosi, E. (2016). *In context: Understanding police killings of unarmed citizen*. St. Augustine, FL: Contextual Press.

Shjarback, J. A., & Nix, J. (2020). Considering violence against the police by citizen race/ethnicity to contextualize representation in officer-involved shootings. *Journal of Criminal Justice*, *66*, 101653. doi:10.1016/j.jcrimjus.2019.101653

Smith, M. R., Rojek, J. J., Petrocelli, M., & Withrow, B. (2017). Measuring disparities in police activities: A state of the art review. *Policing: An International Journal of Police Strategies & Management*, *40*(2), 166–183. doi:10.1108/PIJPSM-06-2016-0074

Smith, M. R., Tillyer, R., Lloyd, C., & Petrocelli, M. (2019). Benchmarking disparities in police stops: A comparative application of 2nd and 3rd generation techniques. *Justice Quarterly*, 1–24. https://www.tandfonline.com/doi/full/10.1080/07418825.2019.1660395. doi:10.1080/07418825.2019.1660395

Streeter, S. (2019). Lethal force in black and white: Assessing racial disparities in the circumstances of police killings. *The Journal of Politics*, *81*(3), 1124–1132. doi:10.1086/703541

Swaine, J., & McCarthy, C. (January 8, (2017). Young black men again faced highest rate of U.S. police killings in 2016. The Guardian, Accessed October 3, 2019. https://www.theguardian.com/us-news/2017/jan/08/the-counted-police-killings-2016-young-black-men.

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

Terrill, W. (2001). *Police coercion: Application of the force continuum*. New York, NY: LFB Scholarly.

Tregle, B., Nix, J., & Alpert, G. P. (2019). Disparity does not mean bias: Making sense of observed racial disparities in fatal officer-involved shootings with multiple benchmarks. *Journal of Crime and Justice*, *42*(1), 18–31. doi:10.1080/0735648X.2018.1547269

Weitzer, R., & Tuch, S. A. (2005). Racially biased policing: Determinants of citizen perceptions. *Social Forces*, *83*(3), 1009–1030. doi:10.1353/sof.2005.0050

Wheeler, A. P., Phillips, S. W., Worrall, J. L., & Bishopp, S. A. (2018). What factors influence an officer's decision to shoot? The promise and limitations of using public data. *Justice Research and Policy*, *18*, 48–76.

White, M. (2016). Transactional encounters crisis-driven reform and the potential for a national police deadly force database. *Criminology & Public Policy*, *15*(1), 223–235. doi:10.1111/1745-9133.12180

Winser, M. A., Hinson, J. M., James, S. M., Vila, B. J., Whitney, P., & Van Dongen, H. P. A. (2014). Fatigue during deadly force decision-making: Measuring skin conductance response during simulations. *Sleep-Wake Research in The Netherlands*, *25*, 81–84.

Worrall, J. L., Bishopp, S. A., Zinser, S. C., Wheeler, A. P., & Phillips, S. W. (2018). Exploring bias in police shooting decisions with real shoot/don't shoot cases. *Crime and Delinquency*, *64*(9), 1171–1192. doi:10.1177/0011128718756038

Appendix

Variable	Black & White Suspects	Black/Hispanic = 1; White = 0	Active Aggression Only
Black Suspect	1.14 (0.76, 1.72)	1.39 (0.94, 2.05)	0.78 (0.31, 1.95)
Black Officer	0.56 (0.31, 1.03)	0.62 (0.37, 1.03)	0.90 (0.25, 3.25)
Male Officer	1.45 (0.73, 2.89)	1.59 (0.90, 2.83)	– –
Male Suspect	5.12*** (2.30, 11.40)	5.51*** (2.72, 11.18)	1.52 (0.34, 6.83)
Officer Tenure	0.96 (0.93, 0.99)	0.96** (0.94, 0.99)	0.97 (0.90, 1.05)
Suspect Mentally Unstable	0.28*** (0.17, 0.47)	0.29*** (0.18, 0.45)	0.30 (0.05, 1.66)
Suspect UI Drugs	0.26*** (0.17, 0.38)	0.28*** (0.21, 0.39)	0.59 (0.20, 1.77)
Risky Call	1.46 (0.87, 2.47)	1.12 (0.70, 1.79)	0.68 (0.09, 5.51)
Suspect Active Aggression	0.35** (0.18, 0.69)	0.37*** (0.21, 0.64)	– –
Suspect Weapon Display	72.43*** (30.44, 172.37)	82.98*** (38.76, 177.64)	– –
Constant	0.06 (0.02, 0.18)	0.05 (0.02, 0.13)	0.12 (0.02, 0.78)
Observations	1,491	1,935	324

Notes: Odds ratios reported (95% confidence intervals in parentheses). * $p < .05$; ** $p < .01$; *** $p < .001$. When estimates were restricted to active aggression cases (last column), there were zero observations for female officers who drew their weapons and zero observations for suspect weapon because suspects could be classified (under available reasons for UoF) as actively aggressive or drawing weapons, but not both.