AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Applied

ISSN: 1076-898X

2025, Vol. 31, No. 4, 297–313
https://doi.org/10.1037/xap0000542

Modeling Police Officers' Deadly Force Decisions in an Immersive Shooting Simulator

Timothy J. Pleskac[1], Joseph Cesario[2], David J. Johnson[2], and Glen Gagnon[2]
[1] Department of Psychological and Brain Sciences, Cognitive Science Program, Indiana University
[2] Department of Psychology, Michigan State University

We used an immersive shooting simulator to examine how race, suspect behavior, and policing scenario shape officers' deadly force decisions. Officers (N = 659) from the Milwaukee Police Department responded to dynamic video scenarios using realistic handgun responses. Mistaken shootings of unarmed Black suspects were more likely than of White suspects, but only when the suspects behaved non-antagonistically. Cognitive modeling showed this race effect arose not from an initial bias to shoot but from differences in evidence accumulation once the object was visible. Scenario and suspect behavior had the largest overall influence, shaping decisions by altering initial proclivity to shoot. Further analysis suggested that suspect behavior within specific scenarios may partially explain observed race effects. These findings provide a process-level account of deadly force decisions, integrating real-world complexity with psychological theory, and offer a framework for improving research and training around police use-of-force.

Public Significance Statement
This study used a realistic police shooting simulator and cognitive modeling to examine how race, suspect behavior, and policing scenarios influence officers' decisions to use deadly force. While officers were more likely to mistakenly shoot unarmed Black suspects in certain ambiguous situations, the more substantial and consistent influences on decision making came from the behavior of the suspect and the nature of the policing scenario. These contextual factors influenced how officers processed information and when they chose to act, underscoring the importance of scenario-based training and realistic decision-making environments in reducing fatal errors.

Keywords: racial bias, racial stereotypes, racial disparities, police use of force, diffusion decision model

Supplemental materials: https://doi.org/10.1037/xap0000542.supp

This article was published Online First August 4, 2025.
Yusuke Yamani served as action editor.
Timothy J. Pleskac https://orcid.org/0000-0001-5761-1900
Joseph Cesario https://orcid.org/0000-0002-1892-4485
Timothy J. Pleskac, Joseph Cesario, and David J. Johnson contributed equally to this article; author order was determined by random draw. All materials, deidentified data, and analysis code are available on the anonymous open science framework link at https://osf.io/bgykz/.

This work was supported by National Science Foundation, Directorate for Social, Behavioral and Economic Sciences Grant 1230281 awarded to Joseph Cesario and Grant 1756092 awarded to Joseph Cesario and Timothy J. Pleskac. David J. Johnson received funding as a postdoctoral researcher at the Lab for Applied Social Science Research at the University of Maryland. Funding was provided by the following Michigan State University offices: vice provost for Research and Graduate Studies, College of Social Science, Department of Psychology, University Outreach and Engagement, and Office for Inclusion and Intercultural Initiatives. The study was approved by the Michigan State Institutional Review Board (13-713). The authors thank officers from the Milwaukee Police Department for their participation. This work would not have been possible without support from Chief Edward Flynn, Captain Nicole Davila, Lieutenant Richard Stein, Sergeant Paul Riehle, Sergeant Jeffrey Sunn, Larry Lumb from Saturn Manufacturing, David McFarlane, and Ahptic Film and Digital. The authors also thank the team of research assistants who helped collect data.

Conceptualization: Joseph Cesario, David J. Johnson, and Timothy J. Pleskac; data curation: Timothy J. Pleskac and David J. Johnson; formal analysis: Timothy J. Pleskac and David J. Johnson; funding acquisition: Joseph Cesario and Timothy J. Pleskac; investigation: Joseph Cesario, David J. Johnson, and Glen Gagnon; methodology: Joseph Cesario, David J. Johnson, and Timothy J. Pleskac; project administration: Joseph Cesario, David J. Johnson, Glen Gagnon, Timothy J. Pleskac: writing–original draft: Joseph Cesario, David J. Johnson. and Timothy J. Pleskac; software: Joseph Cesario; supervision: Joseph Cesario; Timothy J. Pleskac: writing–reviewing and editing: all authors.

Timothy J. Pleskac played an equal role in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, writing–original draft, and writing–review and editing. Joseph Cesario played an equal role in conceptualization, data curation, funding acquisition, investigation, methodology, project administration, software, supervision, writing–original draft, and writing–review and editing. David J. Johnson played a supporting role in writing–original draft and writing–review and editing and an equal role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, and software. Glen Gagnon played a supporting role in investigation, project administration, and writing–review and editing.

Correspondence concerning this article should be addressed to Timothy J. Pleskac, Department of Psychological and Brain Sciences, Cognitive Science Program, Indiana University, 1101 East Tenth Street, Bloomington, IN 47405, United States, or Joseph Cesario, Department of Psychology, Michigan State University, 316 Physics Road, East Lansing, MI 48824, United States. Email: tpleskac@iu.edu or cesario@msu.edu

Police use of deadly force remains a pressing issue in the United States, raising critical questions about how officers make the decision to shoot and how factors such as race influence that decision. These questions have drawn sustained interest across multiple disciplines, particularly psychology and criminology, each of which has brought distinct methodological approaches to bear on the topic. Laboratory-based studies from psychology provide experimental control to isolate psychological mechanisms, while archival analyses from criminology of actual shootings offer realism and ecological validity. Yet both approaches face important limitations that constrain what can be learned about officers' decision-making processes. Here, we develop a third approach: An immersive shooting simulator (ISS) designed to preserve the realism of actual policing scenarios while maintaining the experimental control necessary to model the underlying cognitive processes that guide the decision to shoot.

Controlled Experiments: High Internal Validity and Limited Context

The laboratory-based approach, rooted primarily in psychology, employs controlled tasks to assess racial bias in the decision to shoot (Cesario, 2022; Mekawi & Bresin, 2015; Payne & Correll, 2020). In these tasks, participants typically view static images on a computer screen of Black or White men holding either guns or harmless objects. If the suspect is holding a gun, participants are instructed to press a button labeled "Shoot"; otherwise, they press "Don't Shoot." These decisions are made across many trials. The high degree of control in these tasks allows researchers to isolate how race impacts the decision process, with the prevailing explanation being that automatic activation of the Black-violent stereotype makes people more likely or quicker to "Shoot" in response to Black targets. The most consistent finding is that civilian participants (undergraduates or untrained community members) "Shoot" armed Black suspects faster than White suspects and "Don't Shoot" unarmed Black suspects slower than White suspects (Correll et al., 2002; Mekawi & Bresin, 2015). Some studies also report higher false alarm rates for Black suspects under time pressure (Correll, Park, Judd, & Wittenbrink, 2007; Correll, Park, Judd, Wittenbrink, et al., 2007; Correll et al., 2002, 2011; Pleskac et al., 2018). However, a meta-analysis found the average effect on shooting unarmed suspects (false alarms) across studies to be near zero (Mekawi & Bresin, 2015). Trained officers, by contrast, tend to exhibit only response time differences or no observable racial bias in shooting behavior (Cesario & Carrillo, 2024; Cox et al., 2014; Johnson et al., 2018; Plant & Peruche, 2005; Sadler et al., 2012; Sim et al., 2013).

A key limitation of this line of research may lie in the way the task is framed. By focusing narrowly on whether officers are more likely to shoot unarmed Black men, the experimental paradigm centers on manipulating race, often at the expense of including the situational and interpersonal information officers typically use to guide their decisions (Cesario, 2022; Cox et al., 2014). The removal of this context may limit the generalizability of findings to real-world use-of-force decisions, a limitation that is particularly consequential when interpreting the effects of race.

This methodological choice reflects, in part, the theoretical background of psychologists studying these decisions. Social psychology has long investigated how race and other social categories influence judgment and behavior (e.g., Bargh, 1999; Duncan, 1976). Theories in this tradition argue that accessible social-category information is most likely to shape decisions in situations that are ambiguous or lack clear, diagnostic cues (Higgins et al., 1977; Jussim et al., 2015; Kunda & Thagard, 1996). In line with this perspective, laboratory-based shooting tasks are typically designed to impose severe time pressure and minimize individuating information. For example, participants rarely receive prior contextual cues such as dispatch information, even though such information is present in most real-world police shootings (Johnson et al., 2018).

Archival Analyses: Realism With Limited Process Insight

An alternative to the experimental approach is to study deadly force decisions by analyzing actual officer-involved shootings. This archival approach has been primarily pursued by researchers in criminal justice. In contrast to experimental studies, this approach examines the impact of multiple factors on officers' decisions, including the nature of the policing scenario and suspect behavior. Findings from this work indicate that (a) situational factors—such as the risk level of the encounter or the suspect's aggressiveness—are strongly associated with officers' decisions to shoot and (b) racial disparities in fatal shootings decrease once these situational factors are taken into account (Cesario et al., 2019; Fryer, 2019; Goff et al., 2016; Klinger et al., 2016; Mentch, 2020; Nix et al., 2017; Ross et al., 2021; Scott et al., 2017; Tregle et al., 2019; Wheeler et al., 2017; Worrall et al., 2018).

However, the archival approach is not without limitations. First, it offers no direct means of studying the cognitive dynamics underlying an officer's decision to use deadly force. The lack of experimental control and the infrequency of shootings make such investigations infeasible. Second, although this approach allows for estimating the role of race after adjusting for contextual factors such as suspect behavior, demographics, or criminal activity (e.g., Cesario et al., 2019; Fryer, 2019; Scott et al., 2017), the uniqueness of each shooting complicates interpretation. A factor in one incident may have a different meaning in another, making superficially similar events qualitatively distinct. As a result, even when racial disparities are observed in a specific type of encounter (e.g., warrant service), it is challenging to generalize such findings. Combined with substantial variation in shooting rates across jurisdictions (Ross, 2015) and the rarity of firearm discharges, this approach imposes significant constraints on what can be learned about the underlying decision-making process.

As with the laboratory approach, the methods used in archival research are shaped by theoretical tradition. Nearly all archival research originates from criminal justice and criminology, fields historically focused on understanding the causes of criminal behavior and the consequences of law enforcement strategies. These core questions naturally orient researchers more toward external factors influencing shootings—such as suspect behavior and policing context—and less toward internal cognitive mechanisms like officer bias. Accordingly, suspect behavior and situational context have been the primary focus of archival work on police shootings (e.g., Binder & Scharf, 1980; Klinger et al., 2016; Selby et al., 2016).

ISS: Realistic Simulations With Cognitive Modeling

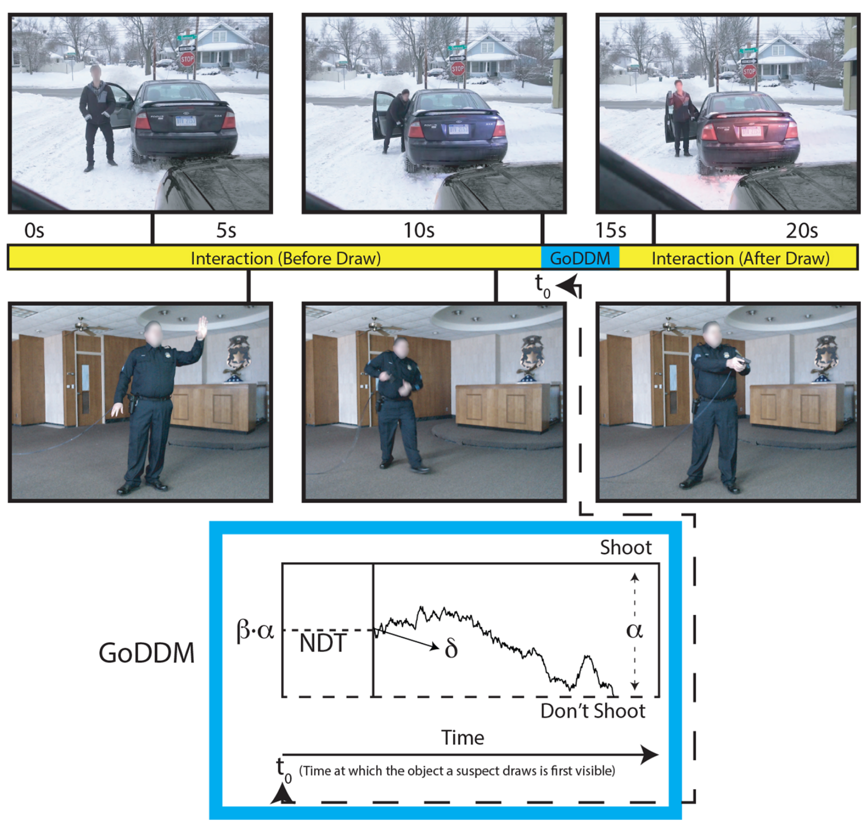
Together, the limitations of the experimental and archival approaches reveal a bottleneck in understanding officer-involved shootings. Laboratory tasks offer controlled tests of specific questions but may lack realism and contextual richness. Archival studies emphasize situational variables but cannot specify how those factors influence cognitive decision making. Neither approach, on its own, is well-suited to informing training interventions aimed at reducing fatal shootings or racial disparities.

To address this gap, we recruited a large sample of officers to make decisions in a shooting simulator modeled on those used in law enforcement training (Figure 1). We developed the ISS to study officers' decisions regarding the use of deadly force in a repeatable and controlled environment that preserves the critical features of policing scenarios. Officers interacted with suspects in life-sized video scenarios filmed from a first-person perspective, following standard protocol. If they decided to use deadly force, they fired a modified handgun that produced realistic sound and recoil and recorded response times. Scenarios were developed in collaboration with law enforcement and depicted a range of encounters (e.g., traffic stops, arrest warrants).

To be sure, a handful of prior studies have used ISSs (Cox et al., 2014; James et al., 2013, 2014, 2016; Taylor, 2020). These studies retained several real-world features—such as dispatch information—that are typically absent from laboratory tasks. They also raised important concerns about the generalizability of lab-based findings to more complex settings. Notably, these immersive studies have not found a race effect (i.e., faster to shoot Black suspects and/or greater likelihood of shooting unarmed Black suspects). Instead, trained officers often show no racial differences (Cox et al., 2014) or are more likely to shoot unarmed White suspects (James et al., 2013, 2016). Some studies have also found that officers are slower to shoot Black versus White suspects (James et al., 2013, 2016).

Nevertheless, these earlier studies have limitations. Some used simplified stimuli even within immersive simulators (Cox et al., 2014), raising familiar concerns about limited individuating information. Others aimed for realism but did not fully cross suspect race and scenario type (James et al., 2013, 2016), making it difficult to isolate the role of race. For example, slower responses to Black

Figure 1
A Trial From the ISS



Note. In this scenario, the officer pulls over a suspect for speeding. The officer talks to the suspect until he reaches into his car, at which time the officer draws his weapon. At approximately 13 s, the critical object (a cell phone) becomes visible for the first time (t_0). The shooting decision process is modeled from the point the object is drawn using the GoDDM (pictured at the bottom) until the decision is made. In this case, the suspect drew a cell phone and the officer did not shoot. ISS = immersive shooting simulator; NDT = nondecision time; GoDDM = go drift diffusion model. See the online article for the color version of this figure.

suspects could reflect the specific scenarios (e.g., hostage situations) in which Black actors appeared.[1] Still others ran a large number of participants (N = 306), but only collected a single trial for each participant (Taylor, 2020). Such an approach is limited in its ability to understand and control the variability inherent in human behavior.

We sought to advance beyond these prior studies by fully crossing suspect race with policing scenario and modeling variation in both scenario and suspect behavior. We conducted this study with a large sample of $N = 659$ police officers who completed up to 32 trials in an immersive simulator, thereby involving a significantly larger number of participants and trials than previous work. Critically, we also extend prior work by applying a computational model to quantify the cognitive processes officers use when deciding to shoot. We now turn to that modeling approach.

Modeling the Cognitive Dynamics of the Decision to Shoot

A formal cognitive model uses mathematical language to specify how basic cognitive processes give rise to observable behavior (Busemeyer & Diederich, 2010; Farrell & Lewandowsky, 2018). In our case, we used the diffusion decision model (DDM; Ratcliff et al., 2016) to characterize officers' decisions to shoot. The DDM models decision making as a process of evidence accumulation toward one of two thresholds, with four psychological parameters capturing core aspects of the decision process (Table 1): the relative start point (β), drift rate (δ), threshold separation (α), and nondecision time (*NDT′*).

Table 1
Parameters of the Go Diffusion Decision Model and Their Substantive Interpretations

Drift diffusion model parameter	Description
Relative start point (β)	The location of the starting point for evidence accumulation relative to the thresholds, with $0 < \beta < 1$. The relative start point indexes an initial bias for either response, with values of β greater than .5 indicating a bias to choose "Shoot" and values less than .5 indicating a bias to not shoot.
Threshold separation (α)	The separation between the thresholds, with $0 < \alpha$. With this parameterization, the choice threshold for "Shoot" is set at α, and the choice threshold for the "Don't Shoot" (unobserved) option is set at 0. The threshold separation determines how much a person trades accuracy for speed (i.e., the speed–accuracy tradeoff), with larger values indicating more accurate but slower decisions.
Drift rate (δ)	The average strength in evidence at each unit of time, with $-\infty < \delta < \infty$. The sign of the drift rate indicates the average direction of the incoming evidence, with negative values indicating evidence in favor of "Don't Shoot" and positive values indicating evidence in favor of "Shoot." The magnitude of the drift rate characterizes the quality of the sampled information.
Relative NDT′	Proportion of response time (relative to the minimum observed response time) spent on processes unrelated to decision making, with $0 < \text{NDT}' < 1$. The nondecision time includes the time spent on encoding the stimulus, executing a response, and any other contaminant process.

Note. NDT = nondecision time.

The DDM is fit to each participant's trial-level data, specifically, their observed choices (i.e., whether they shot or not) and associated response times (when a shot was made). To fit the model, we use a Bayesian hierarchical approach (Lee & Wagenmakers, 2014; Kruschke, 2014), which pools information across participants while allowing for individual-level variation in the parameters. This approach estimates psychologically interpretable parameters for each individual: the relative starting point (β), drift rate (δ), threshold separation (α), and *NDT′*. These parameters have been validated against both cognitive (e.g., Voss et al., 2004) and, to some extent, neural data (see review in Gold & Shadlen, 2007). The DDM has proven effective in describing decisions to shoot in prior laboratory-based research (Correll et al., 2015; Johnson et al., 2018; Pleskac et al., 2018).

A unique aspect of our task is that officers make only one explicit response: to shoot. In contrast, the standard DDM is designed for two-choice tasks where both decisions (e.g., "Shoot" and "Don't Shoot") are observed with associated response times. However, research has shown that go/no-go decisions can still be modeled using the DDM framework, with the "no-go" response treated as an implicit boundary crossing (Gomez et al., 2007; Ratcliff et al., 2018). In our case, we treat "Don't Shoot" responses as censored data—responses that occurred internally but for which we do not observe a time. Our Bayesian estimation technique enables us to jointly model the probability of not shooting and infer the corresponding decision parameters, accounting for both observed and missing components.

We refer to this adaptation as the GoDDM. Figure 1 illustrates the model structure and Table 1 summarizes the parameter interpretations. Full modeling details, including prior specifications and model validation procedures, are provided in the Supplemental Material.

Hypotheses and Questions

Our immersive simulator, combined with the GoDDM cognitive model, enabled us to examine three major sources of variability in officers' deadly force decisions: suspect race, suspect behavior, and the nature of the policing scenario. For each, we evaluated both behavioral outcomes (i.e., errors and response times) and cognitive-level mechanisms (i.e., the model parameters). This dual approach allowed us to test different predictions drawn from psychology and criminal justice, as well as to explore when and how race effects might emerge under more realistic and context-rich conditions.

Suspect Race

Prior laboratory work with untrained civilians has consistently found race effects: participants shoot armed Black suspects faster—and in some cases, more frequently—than armed White suspects (Correll, Park, Judd, Wittenbrink, et al., 2007; Correll et al., 2002). Among police officers, however, race effects are typically limited to faster response times for armed Black suspects, with no differences in error rates (Johnson et al., 2018). Meanwhile, results from immersive simulator studies are mixed, with some reporting null effects

[1] James et al. (2018) also reported work on escalation using simulations that cross race and scenario, though this work does not directly address shooting decisions but instead escalation versus deescalation judgments.

(Cox et al., 2014) and others showing greater errors for White suspects or slower response times for Black suspects (James et al., 2013, 2016). Interpretation of these simulator-based findings is challenging, as past studies often employed limited stimuli or failed to fully cross-race with scenario type.

Our design addresses these limitations. At the behavioral level, we test whether race effects emerge in terms of errors, response times, or both when suspect race and scenario are fully crossed. At the cognitive level, the GoDDM allows us to ask how race may enter the decision process. Based on prior work (Pleskac et al., 2018), two pathways are possible: (1) through the relative start point (β), where officers begin the decision process closer to the "Shoot" threshold for Black suspects (i.e., a "trigger-happy" bias) or (2) through the drift rate (δ), reflecting differences in the evidence accumulated from the scene depending on the suspect's race. For example, officers may accumulate stronger evidence to shoot armed Black suspects and weaker evidence to withhold fire from unarmed Black suspects, relative to White counterparts.

In prior lab-based work, civilians showed race effects in drift rate, while police officers showed race effects in the starting point (Johnson et al., 2018). However, we found that adding even minimal context (e.g., dispatch information) eliminated race effects at both levels. This raises a critical question: will more immersive, naturalistic conditions—like those in our study—similarly attenuate race effects, or might they reemerge when behavior or scenario context is ambiguous?

Suspect Behavior

Both psychological and criminal justice literatures suggest that aggressive suspect behavior increases perceived threat and the likelihood of a shooting. Social psychological theory also predicts that racial stereotypes are more likely to influence decisions under ambiguous conditions, such as when suspects are compliant, because individuating cues are weaker compared with when a suspect is behaving in a clearly aggressive manner. This aligns with broader concerns about fatal shootings involving unarmed individuals who are not acting aggressively (e.g., Philando Castile; Domonoske & Chappell, 2016; Pheifer & Peck, 2016).

To test these predictions, we systematically manipulated suspect behavior across scenarios: suspects portrayed either antagonistic or compliant behavior. At the cognitive level, we examined how suspect behavior impacts model parameters, particularly how it might moderate any race effects on evidence accumulation. If race serves as a cue primarily under ambiguity, we expect differences in drift rate between Black and White suspects to be more pronounced under compliant (nonantagonistic) conditions.

Policing Scenario

Finally, we explored how variations in the policing scenario influence the decision to shoot. While rarely considered in psychological studies, scenario-level features are a central focus in criminal justice research. Officers' threat assessments—and their decision to use deadly force—are strongly shaped by contextual cues such as why they were dispatched or what risks the situation appears to involve.

Because we fully crossed suspect behavior and race across a range of common policing scenarios (e.g., traffic stops, warrant service), we were able to use the GoDDM to quantify how scenario features affect decision processes. In particular, we tested whether certain scenarios shifted officers' initial bias (start point), altered their caution (threshold separation), or affected how quickly and strongly they accumulated evidence toward "Shoot" versus "Don't Shoot."

In sum, our design enables us to test whether race effects emerge at the behavioral level, how they are shaped by suspect behavior, and how scenario context influences the decision to shoot. Across these questions, we apply cognitive modeling to uncover the underlying decision processes that give rise to observed patterns. Together, these analyses provide a more nuanced understanding of when, where, and how race influences officers' decisions to use deadly force.

Method

Participants

Sworn officers ($N = 659$) from the Milwaukee Police Department participated in the study during the spring of 2017. Recruitment occurred between February 27 and May 4, 2017, during a scheduled in-service period. Because there were no prior data on which to base an effect size estimate for officers completing immersive simulators with video scenarios, we did not conduct an a priori power analysis. Instead, we recruited as many participants as possible during the predetermined time frame. Consistent with this approach, we adopted an estimation framework for analysis, focusing on effect sizes and reporting Bayesian credible intervals.

Based on visual observation, 592 officers (90%) were men and 67 (10%) were women. We also visually classified 484 officers (73%) as White, 103 (16%) as Black, 55 (8%) as Hispanic, 14 (2%) as Asian, and 3 (1%) as members of other racial or ethnic groups. The sample was broadly representative of the department, although White officers were overrepresented (73% in the sample compared with 63% in the department). In informal conversations, 94% of participating officers reported an average of 11 years of service as a sworn officer ($SD = 7$; range = 0–25 years).

Each day, we introduced the project at in-service sessions as a study on expert decision making, focusing on fast shoot/do not shoot decisions. We explained that participants would complete multiple scenarios but did not mention that race was a variable of interest. After this introduction, officers voluntarily signed up for individual time slots scheduled throughout the day. No compensation was provided. Data collection spanned a 10-week period during the department's spring in-service. At the time, the department employed approximately 1,800 sworn officers, though not all had the opportunity to participate, as the research team was not present every day.

Officers completed the task individually and at their own pace, within a 20-min time slot. Each officer viewed up to 32 scenarios. Most officers (75%) completed all trials, and nearly all (89%) completed at least 24 trials. Some completed fewer trials due to time constraints or technical difficulties. We excluded trials in which officers responded before the object was first visible ($n = 60$ trials), responses more than three standard deviations above the scenario-specific mean response time ($n = 127$ trials), or trials in which the weapon malfunctioned ($n = 97$ trials). In total, we excluded 284 of 19,600 trials (1.44%). All analyses were conducted on the final sample of 19,316 trials. Analyses using the full data set produced the same conclusions; full results are available on the open science framework (OSF) site at https://osf.io/bgykz/.

Procedure

Officers were informed that the research aimed to understand how experts make fast decisions, particularly regarding object identification and the decision to shoot. They were told they would watch a series of policing scenarios and were instructed to interact with suspects as they would on the job. Officers were informed that if a suspect in the video pulled a gun, the suspect would fire at the officer, and the officer should fire the modified handgun in response. Officers provided consent to have their data used for research purposes and were then asked to consent to having their session filmed.

Each trial began with officers holstering their weapons and reading dispatch information displayed on the screen at their own pace. When they indicated they had read the information, the trial commenced. After each trial, officers reholstered their guns, and dispatch information for the next scenario was presented. Officers were thanked and dismissed after completing all trials or the allotted 20 min.

Materials

Shooting Simulator

Scenarios were displayed using a custom-built shooting simulator similar to commercial law enforcement training simulators. Videos were projected at near life-size. Officers began each trial approximately 15 feet from the screen and were encouraged to speak to suspects and move around as needed. Shooting responses were made using a Glock handgun modified with a Dvorak Air Recoil System. This system replaced the magazine and barrel with a compressed CO_2 system, which cycled the gun and provided recoil when the trigger was pulled.

We further modified the system so that each trigger pull physically activated a microswitch, signaling the computer that the trigger had been pulled; this provided near-millisecond-level timing accuracy.[2] The signal triggered the computer to play the sound of a Glock handgun firing a live round through speakers placed near the screen. All aspects of video presentation and response recording were controlled using PsychoPy (Peirce et al., 2019). Detailed plans are available on the OSF page (https://osf.io/bgykz/).

Video Scenarios

We collaborated with the Milwaukee Police Department to design and film a set of realistic scenarios commonly encountered by officers. We filmed eight scenarios. A full description of each scenario is in the Supplemental Materials, including a table that shows how the different actors were manipulated across the scenarios (Supplemental Table S1). Scenarios were filmed from the officer's point of view and lasted approximately 20 s. All scenarios followed a similar structure. After an initial interaction with a suspect, there were two pivotal moments: one in which the suspect performed an ambiguous action that raised the perceived threat level (e.g., reaching into a glove box), and another in which the suspect drew either a harmless object or a firearm. It was at this second moment that officers had to decide whether to shoot. If the suspect drew a firearm, they always fired at the officer. Officers were under time pressure—the suspect always fired within 1 s of drawing the weapon. Although draw times varied across scenarios, within each scenario, draw time was digitally synchronized across suspect races to within a single video frame.[3]

We employed 10 Black male actors and 10 White male actors as suspects. Each actor was filmed twice per scenario (across at most two different scenarios). In one video, the actor drew a handgun and fired; in the other, the actor revealed a harmless object (e.g., wallet, cellphone). Within each scenario, actors were matched in age, height, and clothing type. Clothing was nondiagnostic of socioeconomic status. Actors wore jeans, t-shirts, and, where appropriate, jackets. The clothing did not have logos or insignias.

We also manipulated the degree to which suspects escalated the interaction. In each scenario, one version featured an actor behaving antagonistically toward the officer, while the other featured an actor behaving nonantagonistically. For example, in a traffic stop scenario, the antagonistic version had the suspect exit the vehicle while swearing at the officer. In contrast, the nonantagonistic version had the suspect exit with hands raised, pleading for calm. This manipulation was fully crossed with suspect race and armed status within each scenario.

Before each scenario, officers were presented with brief dispatch information explaining why they had been called to the scene. Dispatch information was randomly varied and blocked within each scenario. In total, we created 64 unique videos from the eight scenarios. Each scenario had an antagonistic and nonantagonistic version, crossed by suspect race (Black, White) and object (gun, nongun). Officers viewed half of the scenarios, which were randomized so that they could not predict whether a suspect was armed based on prior versions they had seen.

Measures

On each trial, we recorded whether the officer fired their weapon, the response time associated with the first shot, and the number of shots fired. For officers who consented to be filmed (95%), we also coded whether and when they reached for their weapon. For this measure, we coded only the first eight trials of each session to reduce the influence of anticipatory effects that might arise from scenario repetition.

We were able to code 90% of the 4,058 filmed trials. The remaining 10% could not be coded due to officers stepping out of frame, insufficient lighting, or technical issues, yielding a final coded sample of 3,656 trials.

Given that prior research using laboratory-based shooting tasks has focused on shooting decisions and associated response times, we

[2] Trigger pulls physically depressed a lever microswitch positioned in the gun's grip. This microswitch was connected to a modified ioLabs button box via audio cables. The button box was further connected to the computer via USB cable, with PsychoPy reading this signal as a button press. Thus, the signal traveled at the speed of signals transmitted through physical cable lines.

[3] The simulator was programed so that if a participant completed all 32 trials, there would be an equal number of trials (4) across the conditions formed by crossing suspect race, object type, and antagonistic behavior. Due to a programming error, the first nine participants had five trials in the Black, unarmed, antagonistic condition.

present full analyses of these measures in the main text. Additional measures and analyses are reported in the Supplemental Materials.

Analytic Approach

Behavioral Modeling

We analyzed decisions using multilevel logistic regression with suspect race (White, Black), object (gun, nongun), behavior (antagonistic, nonantagonistic), and their interactions as fixed effects. Response times for armed targets were analyzed using multilevel linear regression with suspect race and behavior as fixed effects. To account for individual variability, we included random intercepts for officers.

Initial examination of the data revealed substantial heterogeneity in behavior across scenarios and suspects. To determine the optimal model structure, we compared several candidate models using leave-one-out cross-validation (see Supplemental Materials). For error rates, the best-performing model included random intercepts for scenarios and suspects and random slopes for conditions nested within scenarios. For response times, the best-fitting model included random intercepts for each unique scenario by suspect behavior combination (i.e., video). However, a model with random intercepts for scenarios and suspects and random slopes for nested conditions produced nearly indistinguishable results (i.e., within the margin of error) and led to the same substantive conclusions. Therefore, for consistency, we report results from this latter model in the main text.

In the article, we report the credible effects and credible contrasts. Full summaries of the estimated models are provided on the OSF site (https://osf.io/bgykz/), including posterior distributions of regression coefficients, estimates of mean error rates and response times for each condition, and estimates of race effects within object-by-behavior conditions.

All multilevel models were estimated using Markov Chain Monte Carlo methods implemented via the rstanarm package (Goodrich et al., 2018) in R. This approach enables full Bayesian statistical inference via posterior estimation (Kruschke, 2014). We ran four Markov Chain Monte Carlo chains with 9,000 samples per chain (after a 1,000-sample burn-in), ensuring an effective sample size of over 10,000 for each coefficient. We evaluated convergence using visual inspection and the Gelman–Rubin statistic ($\hat{R}$; Gelman & Rubin, 1992). Default weakly informative priors from rstanarm (Version 2.32.1) were used.

We followed a Bayesian estimation framework. For each parameter or statistic of interest, we report the posterior predicted mean and its 95% highest density interval (HDI; Kruschke, 2014). The HDI defines the most credible values from the posterior distribution. We use the term *credible* throughout to describe effects whose 95% HDI excludes the region of practical equivalence (ROPE). Following convention, the ROPE was defined as $\pm 0.1 \times SD$ (statistic), where the statistic corresponds to a null effect. To assess whether a posterior distribution is centered near the ROPE, we report the proportion of the posterior density that falls within it (p_{ROPE}). A high proportion suggests the effect is likely negligible.

Cognitive Modeling

We used hierarchical Bayesian methods to estimate the GoDDM (Kruschke, 2014; Lee & Wagenmakers, 2014; Vandekerckhove et al., 2011). This approach simultaneously models individual- and group-level parameters, making it well-suited for data sets like ours, where a large number of participants each complete a relatively small number of trials. It also allowed us to model the implicit "Don't Shoot" response by imputing missing response time data.

Consistent with the behavioral modeling, we adopted an estimation approach, reporting the posterior predicted mean of each parameter or statistic of interest M along with its 95% HDI in brackets. We refer to an effect as *credible* when the 95% HDI excludes the ROPE. For the GoDDM, ROPE values were defined as follows: for the relative start point β and relative nondecision time NDT', which are proportions, we set the ROPE to ± 0.05; for the threshold separation α and drift rate δ, which are defined in standardized space, we used ± 0.1. As in the behavioral modeling, we also report the proportion of the posterior distribution's density that falls within the ROPE, denoted as p_{ROPE}.

Parameter Recovery. Model recovery analyses confirmed that the Bayesian framework and experimental design allowed for reasonably accurate recovery of GoDDM parameters (see Supplemental Material). Due to the small number of trials per condition, the recovery analyses revealed a slight underestimation of the relative start point ($\sim$.03) and a slight overestimation of the threshold separation ($\sim$.05). In addition, the between-participant variability parameter for the relative start point was poorly recovered. Some of this imprecision, especially for threshold separation, likely stemmed from using a random-walk approximation of the diffusion process in simulation; the rest likely reflected limited trial counts at the individual level.

We also used the recovery analysis to test whether we could detect effects of race when they existed in only a subset of conditions (e.g., nonantagonistic trials). We simulated two models: one with a race effect in the relative start point and another with a race effect in the drift rate. In both, we coded the effect to be approximately 0.75 in standardized mean difference (in terms of between-participant standard deviation). In each case, we successfully recovered the effect with high accuracy. Overall, the parameter recovery—particularly for between-condition effects—demonstrated that both the direction and magnitude of effects were preserved (i.e., low Type S and Type M error rates; Gelman & Carlin, 2014; Gelman & Tuerlinckx, 2000). Thus, even when a credible effect is not found, the parameter estimates themselves remain informative.

Model Comparison. We parameterized the GoDDM to investigate how object type, suspect race, and suspect behavior influenced each model parameter (see Table 1). We began with a base model in which the start point and threshold varied by race, and the drift rate and nondecision time varied by both race and object. All parameters included random intercepts at the participant level. This base structure reflects prior GoDDM models used in lab-based first-person shooter tasks (Johnson et al., 2019; Pleskac et al., 2018).

Our behavioral results highlighted the importance of accounting for heterogeneity due to differences in scenarios and suspects. However, model recovery analyses showed that it was not possible to incorporate scenario and suspect effects into all four parameters and still reliably recover parameter values. To address this, we conducted model comparisons to determine which parameter best captured these situational effects.

The full comparison is detailed in the Supplemental Material. Briefly, we compared multiple GoDDM models that isolated scenario and suspect effects to a single parameter at a time, using Deviance Information Criterion as our model selection metric (Spiegelhalter et al., 2002). Deviance Information Criterion is particularly suited to hierarchical models and penalizes model complexity. The best-fitting model indicated that both scenario and suspect effects were best captured through variation in the relative start point. A more complex model that included variability in the start point for scenarios, suspects, and officers provided a slightly worse deviance information criterion fit than the model with suspect-only variability. Nonetheless, we chose to use this full model for all reported analyses, as it aligned with our behavioral modeling strategy and reflected our a priori belief that scenarios contribute meaningfully to decision variability.

Transparency and Openness

We are committed to reproducible science. This study was not preregistered. We report how we determined our sample size, all data exclusions, all manipulations, and all measures used in the study. To promote transparency, we have thoroughly documented our procedures and analyses.

All materials (including video stimuli), deidentified data, and analysis code are available on the following OSF repository at https://osf.io/bgykz/ (Pleskac et al., 2025). Supplemental Materials are included with the article. An anonymous review link to the OSF repository (https://osf.io/bgykz/) is provided in the author note.

Results

We analyzed our data at both the behavioral and cognitive levels. Behavioral analyses focused on error rates and response times for armed suspects. Cognitive analyses used the GoDDM to estimate process-level parameters that underlie observed choices and response times.

Behavioral Analyses

We began by analyzing the behavioral data on errors and response times in the decision to shoot. These outcomes were modeled using Bayesian hierarchical models. To account for individual differences, we included random intercepts for officers. Model comparisons (see *Analytic Approach* and Supplemental Materials) revealed that the best-performing models included random intercepts for scenarios and suspects, as well as random slopes for conditions nested within scenarios.

Error Rates

Figure 2A shows error rates by suspect race, object type (weapon presence), and suspect behavior. There was a credible three-way interaction among suspect race, weapon presence, and suspect behavior ($b = 2.21$, [0.39, 4.04], $p_{\text{ROPE}} = 0$).[4] This interaction was primarily driven by the effect of race on unarmed, nonantagonistic suspects: unarmed, nonantagonistic Black suspects were more likely to be shot than unarmed, nonantagonistic White suspects ($M = 1.52$, [0.13, 2.81], $p_{\text{ROPE}} = 0$). However, there were no credible differences between Black and White suspects in other conditions. Specifically, there was no credible difference between Black and White nonantagonistic armed suspects ($M = 0.27$, [−0.94, 1.49], $p_{\text{ROPE}} = 0.13$). There was also no credible difference between Black and White antagonistic suspects when they were unarmed ($M = -0.69$, [−2.04, 0.61], $p_{\text{ROPE}} = 0.07$) or armed ($M = 0.27$, [−0.78, 1.40], $p_{\text{ROPE}} = 0.14$).

At first glance, this three-way interaction appears consistent with prior findings suggesting that decisions to shoot are influenced by threat perception and that race may serve as a cue to threat under ambiguous conditions (Correll et al., 2011). In contrast, in nonantagonistic conditions, where behavior provides little diagnostic information, racial stereotypes may shape perceived threat. In contrast, when behavior is clearly antagonistic, it is likely to dominate the threat assessment, thereby minimizing any racial bias. However, we did not observe a credible increase in overall error rates between antagonistic ($M = 2.7\%$, [0.7, 5.2]) and nonantagonistic conditions ($M = 3.1\%$, [0.9, 6.0]; $b = 0.137$, [−0.771, 1.017], $p_{\text{ROPE}} = 0.35$). Further exploration suggests that part of the observed interaction may stem from the specific behaviors of certain actors in particular scenarios, rather than a race effect per se. We explore this possibility further below, after first reporting the analyses of response times.

Response Times

Figure 2B displays officers' response times to shoot armed suspects, broken down by suspect race and suspect behavior (nonantagonistic vs. antagonistic). There were no credible main effects of race ($b = -0.071$, [−0.257, 0.117], $p_{\text{ROPE}} = .22$), antagonism ($b = 0.010$, [−0.150, 0.171], $p_{\text{ROPE}} = .35$), or a race-by-antagonism interaction ($b = -0.176$, [−0.352, 0.006], $p_{\text{ROPE}} = .04$) on response times.

Additional Behavioral Measures

We also measured several additional behaviors during the ISS: the number of shots fired, whether the officer reached for their gun, and the time at which they reached. Detailed results are provided in the Supplemental Materials. The only credible effect involving suspect race was found in the number of shots fired.

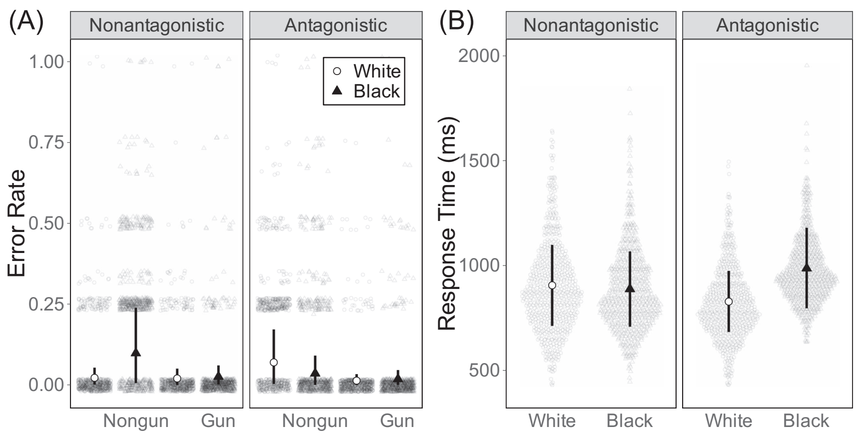
A Poisson regression revealed a credible three-way interaction between suspect race, weapon presence, and suspect behavior ($b = 0.55$, [0.15, 0.94], $p_{\text{ROPE}} = 0$). This interaction was driven by the effect of race on unarmed, nonantagonistic suspects. Officers fired more shots at unarmed, nonantagonistic Black suspects ($M = 2.38$, [1.72, 3.07]) than at unarmed, nonantagonistic White suspects ($M = 1.62$, [1.19, 2.04]; $M_{\text{diff}} = 0.39$, [0.07, 0.68], $p_{\text{ROPE}} = .01$). In contrast, there were no credible race effects for suspects who were armed and nonantagonistic ($M = -0.05$, [−0.22, 0.12], $p_{\text{ROPE}} = .73$), unarmed and antagonistic ($M = -0.14$, [−0.43, 0.15], $p_{\text{ROPE}} = .35$), or armed and antagonistic ($M = -0.03$, [−0.23, 0.16], $p_{\text{ROPE}} = .75$).

We also examined whether any of the behavioral effects were moderated by officers' visually identified gender and race, or their verbally disclosed policing experience. There was no credible four-way interaction between suspect race, weapon presence, suspect behavior, and officer gender ($b = 0.30$, [−1.48, 2.11], $p_{\text{ROPE}} = .15$), officer race ($b = 0.01$, [−1.31, 1.34], $p_{\text{ROPE}} = .22$), or experience ($b = 0.03$, [−0.06, 0.12], $p_{\text{ROPE}} = .22$). Importantly, the three-way

[4] To assess the replicability of this effect, we conducted a retrospective Bayesian power analysis (Kruschke, 2010, 2014; Kruschke & Liddell, 2018). Using the same vague priors (centered at 0, $SD = 2.5$) as in our original analysis, we simulated data from our estimated posterior distributions and tested how often we would recover a credible three-way interaction among suspect race, weapon presence, and suspect behavior. Across 100 simulated data sets, we found a credible three-way interaction in 81% of cases. See the Supplemental Materials for details.

Figure 2

Error Rates and Response Times by Race, Object Type, and Suspect Behavior in the Shooting Simulator



Note. (A) Effect of object and race on errors for nonantagonistic and antagonistic suspects at the group level (solid black outline) and individual level (light gray). (B) Effect of object and race on response times for nonantagonistic and antagonistic suspects at the group level (solid black outline) and individual level (light gray). Nongray markers indicate mean posterior estimates and bars indicate 95% HDIs. All group-level estimates are derived from the group-level posterior predicted distributions from the hierarchical models. HDIs = highest density intervals.

interaction between suspect race, weapon presence, and suspect behavior remained credible when controlling for officer gender and race.

When experience was included as a continuous predictor, the three-way interaction between suspect race, weapon presence, and suspect behavior was not credible (b = 1.93, [−0.22, 4.09], p_{ROPE} = .03), although the effect size remained similar in magnitude. To further explore whether experience moderated this effect, we reran the model using a median split on the experience variable. In this version, the three-way interaction was again credible (b = 2.49, [0.49, 4.49], $p_{\text{ROPE}} = 0$), and the race difference for unarmed, nonantagonistic suspects were similar across both levels of experience. Thus, we conclude that policing experience had minimal impact on the race-related effects observed in the simulator.

Scenario and Suspect Effects

We sought to understand how the interaction between suspect race, object type, and behavior varied across scenarios and individual suspects. Figure 3 plots error rates and response times by scenario for each condition and Figure 4 presents the same data organized by suspect. As these figures show, there was substantial heterogeneity at both the scenario and suspect levels.

Regarding error rates, Figure 3 shows that the effect observed for unarmed, nonantagonistic Black suspects was primarily concentrated in the alley and pullover scenarios. In these videos, the actors drew wallets in a manner that visually mimicked the drawing of a gun. Figure 4 illustrates the corresponding variability across suspects. Unfortunately, our experimental design limits the ability to fully disentangle suspect from scenario, as each suspect appeared in no more than two scenarios.[5]

As noted earlier, real-world decisions to use deadly force are often shaped by situational cues. The ISS enables us to quantify the relationship between variation in the decision to shoot and characteristics of the scenario, the suspect, or the officer. To assess this, we computed intraclass correlations to estimate the proportion of variability in behavior attributable to each of these sources.

For the binary decision to shoot, most of the variance was associated with scenarios (M = 0.19, [0.05, 0.38]) and suspects (M = 0.12, [0.05, 0.22]), with less attributable to officers (M = 0.07, [0.04, 0.09]). In contrast, response time variability was primarily associated with officers (M = 0.19, [0.15, 0.24]) and scenarios (M = 0.14, [0.03, 0.30]), with somewhat less attributable to suspects (M = 0.11, [0.05, 0.19]). These findings indicate that while some officers were consistently faster to shoot than others, decisions about whether to shoot were more strongly influenced by scenario and suspect-level factors.

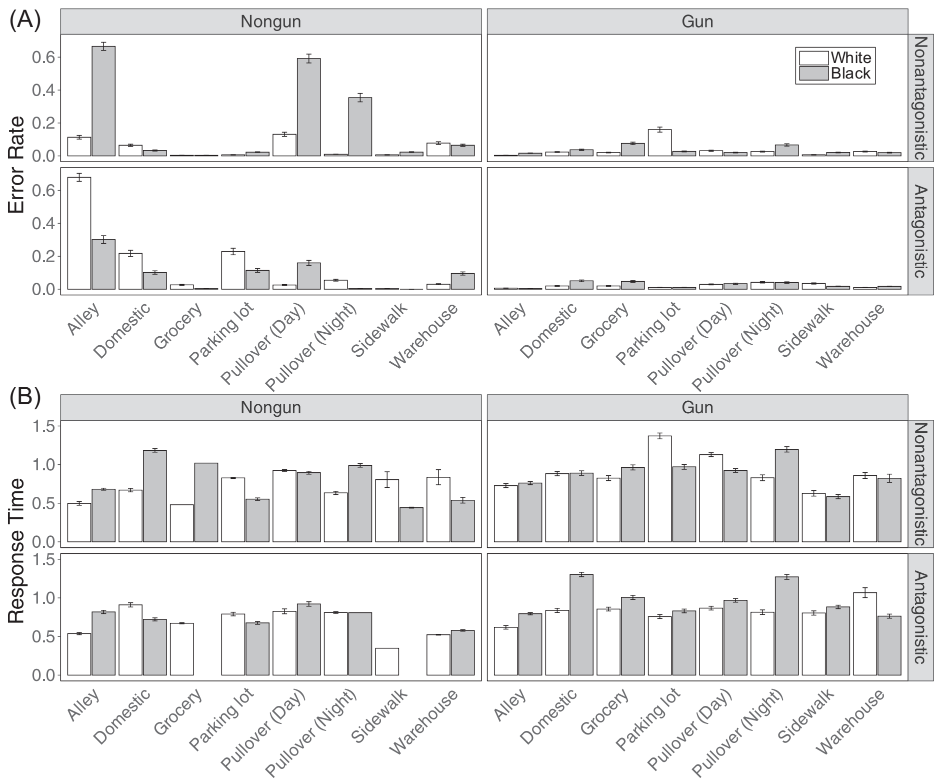
Summary

In sum, the behavioral results demonstrate that multiple features of the decision context shape officers' use-of-force decisions. These features include a policing scenario, suspect behavior, and suspect race. We next turn to cognitive modeling to examine how these factors influenced the underlying decision process, as captured by the parameters of the GoDDM (see Table 1).

[5] For the alley and night pullover scenarios, the Black actors appeared only in those scenarios (suspects 3 and 4 in the alley, suspects 1 and 2 in the pullover). In the alley scenario, the White suspects also appeared only in that scenario.

Figure 3

Error Rates and Response Times by Scenario, Race, Object Type, and Suspect Behavior in the Shooting Simulator



Note. (A) Error rates by scenario, split by whether the suspect is unarmed or armed and whether they acted nonantagonistically or antagonistically. (B) Mean response times by scenario, split by the same factors. Error bars represent 95% confidence intervals.

Cognitive Modeling: GoDDM Analyses

To examine the cognitive processes underlying officers' deadly force decisions, we adapted the Bayesian DDM framework used for the first-person shooter task (Johnson et al., 2018; Pleskac et al., 2018) for the ISS. See Figure 1 for an illustration of the GoDDM process and Table 1 for a description of the model parameters. The decision to shoot is a go/no-go decision where participants must respond to one option (shoot) but withhold their response to another (do not shoot; Donders, 1868/1969). Extensive modeling work has shown that people in a go/no-go procedure use the same evidence accumulation process as when they complete a two-alternative forced-choice task (Gomez et al., 2007; Ratcliff et al., 2018). The difference is that the no-go response (i.e., "Don't Shoot") is implicit (Gomez et al., 2007). That is, participants do decide not to shoot at some point, but its timing is not explicitly known. In our Bayesian framework, we implemented this implicit boundary for the "Don't Shoot" response by treating the response time for this response as missing data. This was done by explicitly modeling the probability that the response time was missing, which in this case was the probability of a "Don't Shoot" response (Kruschke, 2014). Model recovery analyses confirmed that the model parameters can be accurately recovered for this adaptation of the DDM, the GoDDM, with the experimental design used in this study (see Supplemental Material).

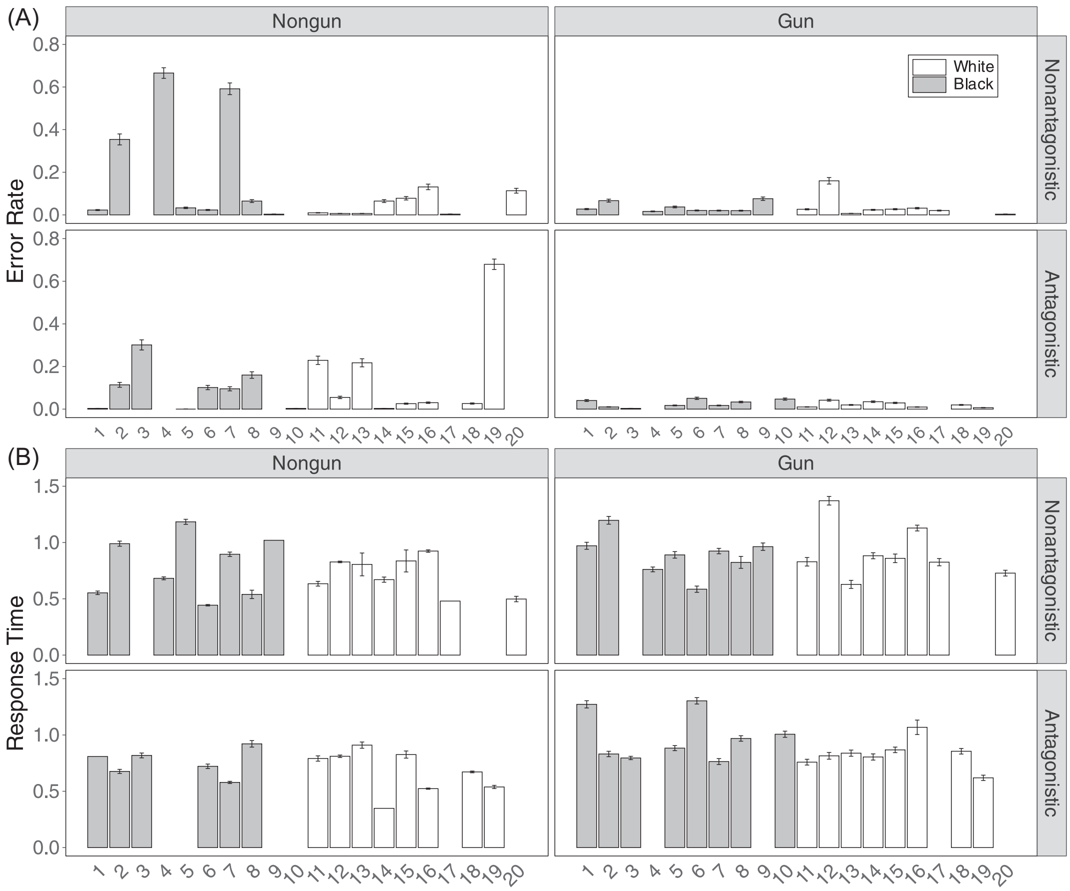
The GoDDM was parameterized to examine how suspect race, suspect behavior, and their interaction impacted each parameter. We modeled participant-level variability by treating participants as a random intercept for each process parameter. Our behavioral analyses highlighted the importance of modeling the heterogeneity in behavior due to differences in the scenario and suspect. Model recovery analyses showed it was impossible to incorporate scenario and suspect effects in each of the four process parameters and accurately recover the values. Therefore, we carried out a model comparison to ask in what process parameter the effect of scenario and suspect should be isolated. The model comparison revealed that the effect of scenario and suspect was in terms of the relative starting point. We use this full model to examine the manipulations of race and suspect behavior on the process parameters.

Process Parameters

Figure 5 plots the group-level estimates for each of the GoDDM parameters, broken down by suspect race and behavior (antagonistic vs. nonantagonistic). A key insight from this cognitive-level analysis is that race and suspect behavior do not exert selective influence on a

Figure 4

Error Rates and Response Times by Suspect, Race, Object Type, and Suspect Behavior in the Shooting Simulator



Note. (A) Error rates by suspect, split by whether the suspect is unarmed or armed and whether they acted non-antagonistically or antagonistically. (B) Mean response times by suspect, split by the same factors. Error bars represent 95% confidence intervals.

single component of the decision process. Rather, these factors give rise to a constellation of effects across multiple parameters.

Relative Start Point. As shown in Figure 5A, there was no credible effect of suspect race on the relative start point β, indicating no initial bias to shoot Black suspects (M = −0.005, [−0.037, 0.028], p_{ROPE} = 0.85). However, there was a credible effect of suspect behavior: officers started farther from the "Shoot" threshold (i.e., closer to "Don't Shoot") when interacting with antagonistic compared with nonantagonistic suspects (M = −0.056, [−0.073, −0.037], p_{ROPE} < .01). This may reflect strategic caution, where officers hesitate to shoot antagonistic suspects who have not yet produced an object.

Importantly, the absence of a race effect in the relative start point suggests that the increased error rate for unarmed, nonantagonistic Black suspects observed at the behavioral level did not stem from a preobject bias to shoot. This finding is inconsistent with the idea that race influenced threat assessment throughout the interaction by shifting officers closer to the "Shoot" threshold before the object appeared.

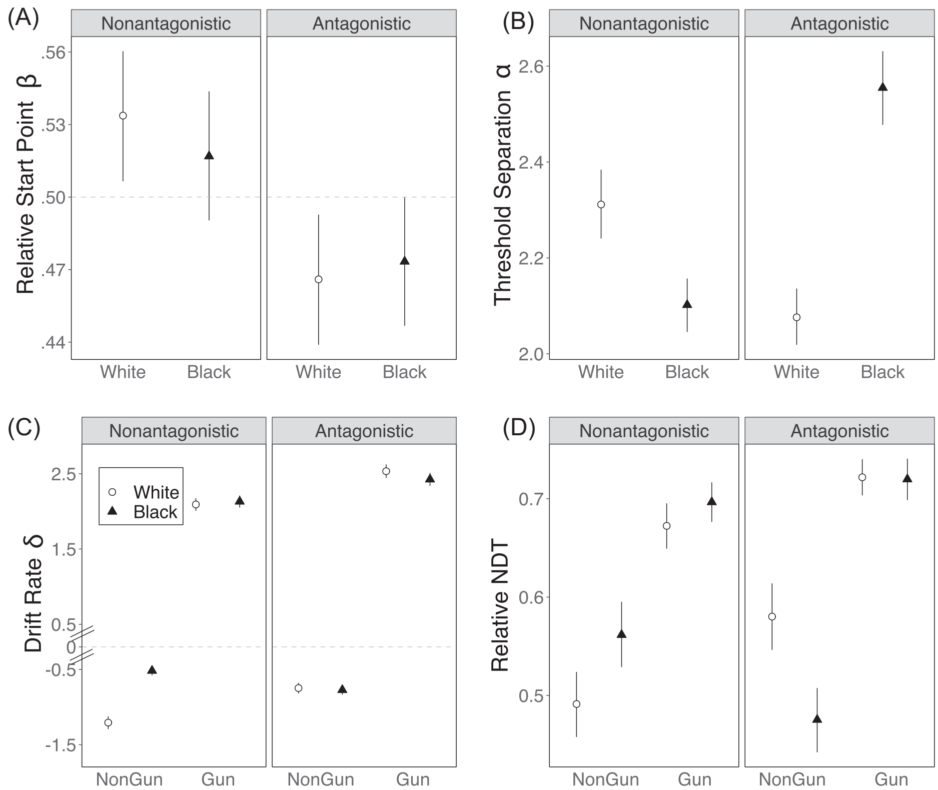
Threshold Separation. Figure 5B displays the threshold separation parameter α, which indexes the amount of evidence required before making a decision and thus reflects caution. Results revealed a credible interaction between race and suspect behavior (M = −0.344, [−0.408, −0.278], p_{ROPE} = 0). Officers exhibited lower thresholds for nonantagonistic Black suspects compared with White suspects (M = −0.210, [−0.298, −0.121], p_{ROPE} < .01), but higher thresholds for antagonistic Black suspects (M = 0.479, [0.385, 0.574], p_{ROPE} = 0). This suggests that officers adjusted their caution based on both the suspect's race and behavior—exercising less caution with Black suspects who were nonantagonistic and more caution with Black suspects who were antagonistic.

At the behavioral level, this pattern may help explain slower response times for armed, antagonistic Black suspects (M = 989 ms, [762, 1,222]) compared with White antagonistic suspects (M = 825 ms, [661, 996]), although this difference was not credible.

Drift Rate. Figure 5C shows posterior estimates for the drift rate δ, which reflects the strength and direction of evidence accumulation. Results showed a credible three-way interaction between suspect race, object type, and behavior (M = −0.216, [−0.276, −0.156], p_{ROPE} = .04). This interaction was primarily driven by differences in the rate of evidence accumulation toward "Don't Shoot" for unarmed,

Figure 5

Go Drift Diffusion Model Group Level Parameters by Race, Object Type, and Suspect Behavior in the Shooting Simulator



Note. (A) Group-level estimates of relative start point by race for nonantagonistic and antagonistic suspects. (B) Group-level estimates of threshold separation by race and behavior. (C) Group-level estimates of drift rate by object type, suspect race, and behavior. (D) Group-level estimates of relative nondecision time by object type, suspect race, and behavior. Markers indicate mean posterior estimates and bars indicate 95% HDIs. HDIs = highest density intervals; NDT = nondecision time.

nonantagonistic suspects: officers accumulated stronger evidence against shooting White suspects than Black suspects in this condition (M = −0.690, [−0.794, −0.587], $p_{\text{ROPE}} = 0$).

No credible race differences were found for drift rates involving unarmed, antagonistic suspects (M = −0.025, [−0.121, −0.069], $p_{\text{ROPE}} = .93$), armed, nonantagonistic suspects (M = 0.039, [−0.073, 0.153], $p_{\text{ROPE}} = .85$), or armed, antagonistic suspects (M = −0.109, [−0.230, 0.014], $p_{\text{ROPE}} = .44$).

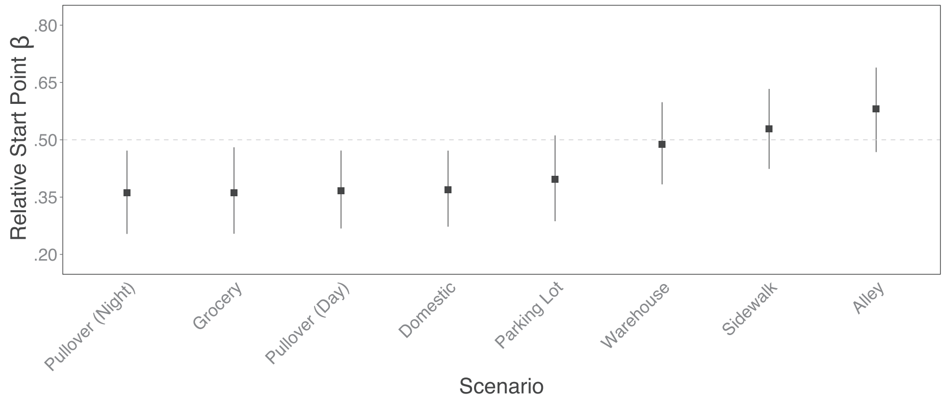
Further comparisons showed that, for White suspects, drift rates toward “Don’t Shoot” were stronger in the nonantagonistic condition than in the antagonistic condition (M = −0.447, [−0.544, −0.352], $p_{\text{ROPE}} = 0$). For Black suspects, the opposite was true—drift rates toward “Don’t Shoot” were weaker in the nonantagonistic condition compared with antagonistic (M = 0.243, [0.167, 0.318], $p_{\text{ROPE}} < .01$). The magnitude of this difference between nonantagonistic and antagonistic conditions was nearly twice as large for White suspects than for Black suspects (M = 1.90, [1.13, 2.77], $p_{\text{ROPE}} < .01$). This is consistent with race functioning as a cue to disambiguate threat in ambiguous (nonantagonistic) contexts. More broadly, the combination of weaker drift rates and lower thresholds for Black nonantagonistic suspects may explain their elevated error rates.

Nondecision Time. Figure 5D summarizes the posterior estimates of the nondecision time parameter NDT', which captures the proportion of time attributed to processes unrelated to deliberation, such as perceptual encoding or motor execution. Officers showed higher nondecision time for unarmed, nonantagonistic Black suspects compared with White suspects (M = 0.07, [0.02, 0.12], $p_{\text{ROPE}} = .03$), but lower nondecision time for antagonistic Black suspects (M = −0.10, [−0.16, −0.06], $p_{\text{ROPE}} < .01$).

Cognitive Modeling: Scenario and Suspect Effects

As described earlier, scenario and suspect effects were modeled as influencing the relative start point. Figure 6 displays the relative start point estimates by scenario. Officers showed a credible bias toward “Shoot” in scenarios where the predecision context suggested elevated threat, such as when suspects had outstanding warrants (e.g., alley and sidewalk scenarios) or where suspects held ambiguous but potentially dangerous objects (e.g., a crowbar in the warehouse scenario). These results map closely onto the behavioral data: officers responded more quickly to armed suspects in the warehouse and sidewalk scenarios and committed more errors in the alley scenario (see Figure 2).

Figure 6
Relative Start Points for Each Scenario



Note. Dots indicate mean posterior estimates and bars indicate 95% HDIs. HDIs = highest density intervals.

Figure 7 displays relative start points for each suspect, separated by antagonistic and nonantagonistic behavior. This figure highlights the substantial variability in initial bias attributable to the suspect, independent of race or behavior. In fact, the between-suspect variability in start point was greater than the variability observed for the other manipulations.

Summary

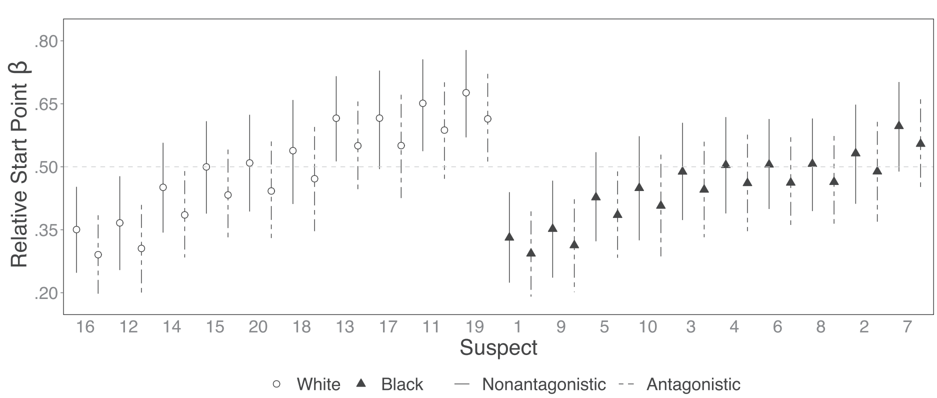
Taken together, the GoDDM results provide insight into how multiple factors influence the decision-making process behind the use of deadly force. When race influenced the decision process, it did not appear to operate through an initial bias to shoot. Instead, it entered the process through evidence accumulation, particularly when suspects behaved in a nonantagonistic manner. Officers also widened their threshold (i.e., became more cautious) when interacting with antagonistic Black suspects, consistent with the strategic caution. This combination of effects explains the increased error rates for unarmed, nonantagonistic Black suspects and slower response times for antagonistic Black suspects.

The GoDDM also captured how situational features—such as the scenario and individual suspect characteristics—influenced decision dynamics. These contextual factors primarily affected the relative start point, reflecting officers' initial proclivity to shoot or withhold fire even before the object was revealed.

Discussion

We utilized an ISS and cognitive modeling to gain insight into officers' decisions regarding the use of deadly force. This unique combination allowed us to investigate how policing scenario, suspect behavior, and race influence the decision-making process. The behavioral analyses support two primary conclusions. First, race influenced shooting decisions in specific conditions: officers were more likely to mistakenly shoot unarmed Black suspects who behaved nonantagonistically. Second, suspect behavior and policing scenario accounted for substantial variability in decisions, often shaping outcomes more than race alone. As we describe next, our computational cognitive model—the GoDDM—isolates the cognitive mechanisms driving these behavioral effects, offering deeper insight into deadly force decision making.

Figure 7
Relative Start Points for Each Suspect



Note. Markers indicate mean posterior estimates and bars indicate 95% HDIs. HDIs = highest density intervals.

Effects of Race on the Decision to Shoot

The ISS and GoDDM framework provide a direct test of hypotheses regarding how race influences the decision-making process. One common hypothesis is that officers have an initial bias to shoot Black suspects. According to the GoDDM, this bias would manifest in officers having a higher starting point (β) closer to "Shoot." Our results do not support this hypothesis. If anything, officers began closer to "Don't Shoot" when interacting with Black suspects. Instead, the behavioral pattern—greater errors for unarmed, nonantagonistic Black suspects—is explained by race influencing the drift rate. Officers showed stronger evidence accumulation toward not shooting for unarmed, nonantagonistic White suspects and toward shooting for Black suspects in the same condition (though still generally accumulating evidence against shooting). At the same time, officers set a lower threshold for evidence for unarmed, nonaggressive Black suspects. These two changes in the cognitive processes explain the increased error rate in this condition.

This difference in drift rates reflects a change in the information being accumulated to decide to shoot. One possible explanation for this difference is that officers may consider factors such as the suspect's race, in addition to the object they are holding, particularly when the suspects are nonaggressive. Stereotypes about Black males as violent may have interfered with object identification when the nongun object was held by a Black suspect. This interpretation aligns with prior work showing that stereotypes shape decisions under ambiguity, but not when behavior provides clear diagnostic cues (Duncan, 1976; Jussim et al., 2015; Kunda & Thagard, 1996; Sagar & Schofield, 1980). These results may parallel findings from the simplified First-Person Shooter Task, where undergraduates showed elevated drift rates toward "Shoot" for Black suspects, particularly when suspects were armed (Johnson et al., 2018; Pleskac et al., 2018). Notably, police officers in that task showed differences in starting point, not drift rate (Johnson et al., 2018). Moreover, for both groups, conditions with minimal context (e.g., dispatch information) yielded no credible race effects for either novice undergraduates or expert police officers. That is, like our results here, race effects were observed only under ambiguous conditions. Future research should seek to better understand the potential role ambiguity plays in the use of racial stereotypes during the decision to shoot, particularly when suspect behavior and context do not provide strong threat cues.

However, these findings must be interpreted in light of notable variability across scenarios and actors that our ISS revealed. For example, the videos with higher error rates featured actors who appeared to draw harmless objects in a manner resembling drawing a gun. Future studies could better isolate these effects through systematic manipulations of object draw behavior and actor characteristics. Additional work using eye tracking or perceptual manipulations could help identify whether difficulties in object recognition or movement style influence the decision to shoot.

Suspect Behavior and Policing Scenario

Our findings also highlight the influence of policing scenarios and suspect behavior on shooting decisions. The ISS enables the study of how these more complex and realistic features influence decision-making processes. At the process level, these factors influenced the initial starting point: officers began the evidence accumulation process closer to "Shoot" or "Don't Shoot" based on situational cues. This links cognitive modeling directly to long-standing insights from both psychology and criminal justice, thereby bridging the methodological divide between these disciplines.

A novel contribution of this study is the ability to disentangle the cognitive impact of race and scenario. Race effects emerged through evidence accumulation (drift rate), while scenario effects emerged through officers' starting bias (relative start point). Importantly, our design orthogonally crossed race and scenario using a systematic experimental approach. In contrast, real-world encounters are not orthogonal: police are more likely to encounter citizens of different racial backgrounds in different types of situations. This correlation raises the possibility that cognitive processes associated with race and context might combine in practice. For instance, for Black citizens, both race-based drift rate effects and scenario-based start point effects may operate simultaneously, while for White citizens, neither may be present. Future studies using less controlled but more ecologically realistic designs could examine how these interactions shape deadly force decisions.

Limitations and Constraints on Generality

While the ISS allows greater ecological validity than standard lab tasks, it remains a simulation. Officers knew they would not die or face formal consequences for incorrect decisions. They were also limited in their response options and could not interact dynamically with suspects in the system (cf. James et al., 2018). Still, officers' behaviors—such as verbal commands and physiological responses—suggest they engaged with the task seriously. The results also align with past research, which shows that situational and suspect-level variables strongly predict use-of-force decisions (Bolger, 2015; National Research Council, 2004; Terrill & Mastrofski, 2002; Terrill & Reisig, 2003; Wheeler et al., 2017; White, 2002).

A related issue is that our study placed officers in a situation where their shooting behavior was being observed. This introduces a potential confound where officers may have adjusted their behavior, especially in interactions with Black suspects. While plausible, the GoDDM results do not support this interpretation. If officers were responding to being observed, we would expect greater vigilance, reflected in increased threshold separation or increased sensitivity (i.e., larger drift rate differences) for Black suspects. Neither pattern was observed. In fact, threshold separation was greatest only for antagonistic Black suspects, where officers would likely feel most justified in using deadly force. Thus, while social desirability effects are possible in our design, the pattern of cognitive model results does not support this explanation.

Another concern raised by our findings is the apparent lack of a main effect of antagonistic behavior on shooting likelihood and response time. While it might be expected that antagonism would increase both the frequency and speed of shootings, our data showed the opposite pattern: officers' initial bias (start point) shifted closer to "Don't Shoot" for antagonistic suspects. We acknowledge this may seem counterintuitive and may reflect, in part, a response strategy shaped by the simulation context. Our design aimed to maximize ecological validity, and, as discussed earlier, several indications in the data suggest that the ISS corresponds well with real-world decision dynamics. Nevertheless, the ISS lacks interaction, real-time escalation, and real-world consequences. All of these aspects may suppress automatic defensive reactions typically evoked by antagonistic behavior in the field. It is possible that officers interpreted antagonism in the ISS not solely as a threat cue, but also as a potential trap, consistent with training experiences designed to assess restraint under pressure. This strategic caution may explain the observed reduction in starting bias to shoot. These findings underscore the importance of exercising caution when generalizing from simulation to real-world behavior and suggest that future studies should investigate how simulation framing and officer expectations impact decision-making strategies in high-stakes contexts.

Indeed, caution is warranted in generalizing these findings. The sample came from a single police department, and departments vary in their policies and practices (Ross, 2015; Terrill et al., 2018). Officers volunteered, and although they were not aware of the study's focus on race, self-selection remains possible (Heckman, 1979). The data were also collected in 2017.[6] The worldwide Black Lives Matter protests (Buchanan et al., 2020; Silverstein, 2021) that followed the killing of George Floyd (McGreal et al., 2021) led to some reform (Lartey, 2023), the consequences of which are not yet well understood. Similarly, we urge caution in generalizing beyond the specific scenarios used in our version of the ISS. As with real-world shootings, each scenario is unique, and features present in one may have different meanings in another. While the ISS enables controlled observation of officers' responses to standardized situations, a richer theory of policing scenarios is needed to generalize these findings more broadly.

Additionally, our focus was specifically on the decision to shoot. These findings may not be generalizable to other forms of force or types of police interactions. Most encounters do not involve armed suspects, and even when they do, not all pose lethal threats. Officers also have options besides deadly force. Our conclusions apply specifically to time-pressured situations where officers must make rapid object identification decisions. Finally, although computational modeling offers powerful insights into latent processes, other methods—such as neural measures (Freeman et al., 2014), eye tracking (Correll et al., 2015), and verbal protocols (Ericsson & Simon, 1984), can further enrich our understanding of these critical judgments.

Conclusion

We used an ISS to examine how race, suspect behavior, and policing scenarios shape officers' deadly force decisions. The ISS bridges the control and precision of laboratory tasks with the ecological realism of field-based variables. While race influenced the decision process in specific, low-threat or ambiguous contexts, the more substantial effects were driven by suspect behavior and policing scenarios. These contextual factors shaped officers' decisions at the cognitive level by altering their starting bias and threshold for action. Together, these findings offer a process-level understanding of deadly force decisions and a framework for developing targeted, data-driven interventions in law enforcement training and policy.

[6] The long time lag between data collection and publication warrants some comment. The public attention to the topic of police killings, particularly during the summer of 2020, pushed us to exercise an abundance of caution in terms of data analysis and our interpretation of the results, both in terms of behavioral and cognitive modeling.

References

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). Guilford Press.

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. https://doi.org/10.1177/000271628045200111

Bolger, P. C. (2015). Just following orders: A meta-analysis of the correlates of American police officer use of force decisions. *American Journal of Criminal Justice*, *40*(3), 466–492. https://doi.org/10.1007/s12103-014-9278-y

Buchanan, L., Bui, Q., & Patel, J. K. (2020, July 3). Black lives matter may be the largest movement in U.S. history. *New York Times*. https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html

Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Sage Publications.

Cesario, J. (2022). What can experimental studies of bias tell us about real-world group disparities? *Behavioral and Brain Sciences*, *45*, Article e66. https://doi.org/10.1017/S0140525X21000017

Cesario, J., & Carrillo, A. (2024). Racial bias in police officer deadly force decisions: What has social cognition learned? In D. Carlston, K. Johnson, & K. Hugenberg (Eds.), *The oxford handbook of social cognition* (2nd ed., pp. 525–559). Oxford University Press.

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. https://doi.org/10.1177/1948550618775108

Chen, Y., Fang, J., Cesario, J., Liu, T., & Pleskac, T. J. (2025). Comparing one-boundary and two-boundary evidence accumulation models for go/no-go processes: An application to the decision to shoot. *Proceedings of the annual meeting of the cognitive science society* (Vol. *47*). Cognitive Science Society.

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. https://doi.org/10.1037/0022-3514.83.6.1314

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007). The influence of stereotypes on decisions to shoot. *European Journal of Social Psychology*, *37*(6), 1102–1117. https://doi.org/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. https://doi.org/10.1037/0022-3514.92.6.1006

Correll, J., Wittenbrink, B., Crawford, M., & Sadler, M. (2015). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology*, *108*(2), 219–233. https://doi.org/10.1037/pspa0000015

Correll, J., Wittenbrink, B., Park, B., Judd, C. M., & Goyle, A. (2011). Dangerous enough: Moderating racial bias with contextual threat cues. *Journal of Experimental Social Psychology*, *47*(1), 184–189. https://doi.org/10.1016/j.jesp.2010.08.017

Cox, W. T., Devine, P. G., Plant, E. A., & Schwartz, L. L. (2014). Toward a comprehensive understanding of officers' shooting decisions: No simple answers to this complex problem. *Basic and Applied Social Psychology*, *36*(4), 356–364. https://doi.org/10.1080/01973533.2014.923312

Diederich, A., & Oswald, P. (2016). Multi-stage sequential sampling models with finite or infinite time horizon and variable boundaries. *Journal of Mathematical Psychology*, *74*, 128–145. https://doi.org/10.1016/j.jmp.2016.02.010

Domonoske, C., & Chappell, B. (2016, July 7). *Minnesota gov. calls traffic stop shooting' absolutely appalling at all levels'*. National Public Radio. https://www.npr.org/sections/thetwo-way/2016/07/07/485066807/police-stop-ends-in-black-mans-death-aftermath-is-livestreamed-online-video

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. https://doi.org/10.1016/0001-6918(69)90065-1 (Original work published 1868)

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology*, *34*(4), 590–598. https://doi.org/10.1037/0022-3514.34.4.590

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. The MIT Press.

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press.

Freeman, J. B., Stolier, R. M., Ingbretsen, Z. A., & Hehman, E. A. (2014). Amygdala responsivity to high-level social information from unseen faces. *Journal of Neuroscience*, *34*(32), 10573–10581. https://doi.org/10.1523/JNEUROSCI.5063-13.2014

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

Gelman, A., & Carlin, J. (2014). Beyond power calculations assessing types (sign) and type m (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., & Tuerlinckx, F. (2000). Type s error rates for classical and bayesian single and multiple comparison procedures. *Computational Statistics*, *15*(3), 373–390. https://doi.org/10.1007/s001800000040

Goff, P. A., Lloyd, T., Gelle, A., Raphael, S., & Glaser, J. (2016). *The science of justice: Race, arrests, and police use of force*. Center for Policing Equity. https://policingequity.org/wp-content/uploads/2020/08/CPE-Report-SOJ-RaceArrestsUseOfForce.pdf

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389–413. https://doi.org/10.1037/0096-3445.136.3.389

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2018). *rstanarm: Bayesian applied regression modeling via Stan* (R package Version 2.17.4) [Computer software]. https://mc-stan.org/

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161. https://doi.org/10.2307/1912352

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*(2), 141–154. https://doi.org/10.1016/S0022-1031(77)80007-3

James, L., James, S. M., & Vila, B. J. (2016). The reverse racism effect: Are cops more hesitant to shoot Black than White suspects? *Criminology & Public Policy*, *15*(2), 457–479. https://doi.org/10.1111/1745-9133.12187

James, L., James, S., & Vila, B. (2018). Testing the impact of citizen characteristics and demeanor on police officer behavior in potentially violent encounters. *Policing: An International Journal*, *41*(1), 24–40. https://doi.org/10.1108/PIJPSM-11-2016-0159

James, L., Klinger, D., & Vila, B. (2014). Racial and ethnic bias in decisions to shoot seen through a stronger lens: Experimental results from high-fidelity laboratory simulations. *Journal of Experimental Criminology*, *10*(3), 323–340. https://doi.org/10.1007/s11292-014-9204-9

James, L., Vila, B., & Daratha, K. (2013). Results from experimental trials testing participant responses to White, Hispanic and Black suspects in high-fidelity deadly force judgment and decision-making simulations. *Journal of Experimental Criminology*, *9*(2), 189–212. https://doi.org/10.1007/s11292-012-9163-y

Johnson, D. J., Cesario, J., & Pleskac, T. J. (2018). How prior information and police experience impact decisions to shoot. *Journal of Personality and Social Psychology*, *115*(4), 601–623. https://doi.org/10.1037/pspa0000130

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. https://doi.org/10.1073/pnas.1903856116

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Jussim, L., Crawford, J. T., Anglin, S. M., Chambers, J. R., Stevens, S. T., & Cohen, F. (2015). Stereotype accuracy: One of the largest and most replicable effects in all of social psychology. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 31–63). Psychology Press.

Karlin, S., & Taylor, H. E. (1981). *A second course in stochastic processes*. Elsevier.

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. https://doi.org/10.1111/1745-9133.12174

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*(7), 293–300. https://doi.org/10.1016/j.tics.2010.05.001

Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, *25*, 178–206. https://doi.org/10.3758/s13423-016-1221-4

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*(2), 284–308. https://doi.org/10.1037/0033-295X.103.2.284

Lartey, J. (2023, June 3). *Three years after George Floyd's murder, police reforms are slow-paced*. CBS News. https://www.themarshallproject.org/2023/06/03/george-floyd-police-reform

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Loughlin, J. K., & Flora, K. (2017). *Shots fired: The misunderstandings, misconceptions, and myths about police shootings*. Simon and Schuster.

McGreal, C., Beckett, L., Laughland, O., & Ajasa, A. (2021, April 21). Derek Chauvin found guilty of murder of George Floyd. *The Guardian*. https://www.theguardian.com/us-news/2021/apr/20/derek-chauvin-verdict-guilty-murder-george-floyd

Mekawi, Y., & Bresin, K. (2015). Is the evidence from racial bias shooting task studies a smoking gun? Results from a meta-analysis. *Journal of Experimental Social Psychology*, *61*, 120–130. https://doi.org/10.1016/j.jesp.2015.08.002

Mentch, L. (2020). On racial disparities in recent fatal police shootings. *Statistics and Public Policy*, *7*(1), 9–18. https://doi.org/10.1080/2330443X.2019.1704330

National Research Council. (2004). *Fairness and effectiveness in policing: The evidence*. National Academies Press.

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 implicit bias. *Criminology & Public Policy*, *16*(1), 309–340. https://doi.org/10.1111/1745-9133.12269

Payne, B. K., & Correll, J. (2020). Race, weapons, and the perception of threat. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 62, pp. 1–50). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.001

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pheifer, P., & Peck, C. (2016, July 7). Aftermath of fatal falcon heights officer-involved shooting captured on video. *Star Tribune*. https://www.startribune.com/aftermath-of-fatal-falcon-heights-officer-involved-shooting-captured-on-video/385789251/

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

Pleskac, T. J., Cesario, J., & Johnson, D. J. (2018). How race affects evidence accumulation during the decision to shoot. *Psychonomic Bulletin & Review*, *25*, 1301–1330. https://doi.org/10.3758/s13423-017-1369-6

Pleskac, T. J., Cesario, J., Johnson, D. J., & Gagnon, G. (2025). *Modeling police officers' deadly force decisions in an immersive shooting simulator*. https://osf.io/bgykz/

Plummer, M. (2003). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling* [Computer software]. https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-go task with a diffusion model. *Decision*, *5*(1), 42–62. https://doi.org/10.1037/dec0000065

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

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), Article e0141854. https://doi.org/10.1371/journal.pone.0141854

Ross, C. T., Winterhalder, B., & McElreath, R. (2021). Racial disparities in police use of deadly force against unarmed individuals persist after appropriately benchmarking shooting data on violent crime rates. *Social Psychological and Personality Science*, *12*(3), 323–332. https://doi.org/10.1177/1948550620916071

Rylant, C. J. (2017). *Shots fired: The psychology behind officer involved shootings*. Perfect Life.

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. https://doi.org/10.1111/j.1540-4560.2012.01749.x

Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in Black and White children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology*, *39*(4), 590–598. https://doi.org/10.1037/0022-3514.39.4.590

Scott, K., Ma, D. S., Sadler, M. S., & Correll, J. (2017). A social scientific approach toward understanding racial disparities in police shooting: Data from the department of justice (1980–2000). *Journal of Social Issues*, *73*(4), 701–722. https://doi.org/10.1111/josi.12243

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

Silverstein, J. (2021, July 4). *The global impact of George Floyd: How Black lives matter protests shaped movements around the world*. CBS News. https://www.cbsnews.com/news/george-floyd-black-lives-matter-impact/

Sim, J. J., Correll, J., & Sadler, M. S. (2013). Understanding police and expert performance: When training attenuates (vs. exacerbates) stereotypic bias in the decision to shoot. *Personality and Social Psychology Bulletin*, *39*(3), 291–304. https://doi.org/10.1177/0146167212473157

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Steyvers, M. (2011). *Matjags 1.3: A matlab interface for jags* [Computer software]. https://github.com/msteyvers/matjags

Taylor, P. L. (2020). Dispatch priming and the police decision to use deadly force. *Police Quarterly*, *23*(3), 311–332. https://doi.org/10.1177/1098611119896653

Terrill, W., & Mastrofski, S. D. (2002). Situational and officer-based determinants of police coercion. *Justice Quarterly*, *19*(2), 215–248. https://doi.org/10.1080/07418820200095221

Terrill, W., Paoline, E. A. III, & Ingram, J. R. (2018). Beyond the Final Report: A research note on the assessing police use of force policy and outcomes project. *Policing: An International Journal*, *41*(2), 194–201. https://doi.org/10.1108/PIJPSM-04-2017-0047

Terrill, W., & Reisig, M. D. (2003). Neighborhood context and police use of force. *Journal of Research in Crime and Delinquency*, *40*(3), 291–321. https://doi.org/10.1177/0022427803253800

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. https://doi.org/10.1080/0735648X.2018.1547269

Tuerlinckx, F., Maris, E., Ratcliff, R., & De Boeck, P. (2001). A comparison of four methods for simulating the diffusion process. *Behavior Research Methods, Instruments, & Computers*, *33*(4), 443–456. https://doi.org/10.3758/BF03195402

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44–62. https://doi.org/10.1037/a0021765

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Wabersich, D., & Vandekerckhove, J. (2014). Extending jags: A tutorial on adding custom distributions to jags (with a diffusion model example). *Behavior Research Methods*, *46*(1), 15–28. https://doi.org/10.3758/s13428-013-0369-3

Wheeler, A. P., Phillips, S. W., Worrall, J. L., & Bishopp, S. A. (2017). What factors influence an officer's decision to shoot? The promise and limitations of using public data. *Justice Research and Policy*, *18*(1), 48–76. https://doi.org/10.1177/1525107118759900

White, M. D. (2002). Identifying situational predictors of police shootings using multivariate analysis. *Policing: An International Journal of Police Strategies & Management*, *25*(4), 726–751. https://doi.org/10.1177/1525107118759900

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 & Delinquency*, *64*(9), 1171–1192. https://doi.org/10.1177/0011128718756038

Received December 5, 2024
Revision received June 10, 2025
Accepted June 16, 2025 ■