MODELING DECISION PROCESSES IN THE USE OF LETHAL FORCE: THE ROLE OF RACIAL BIAS IN JUDGING FACES

By

Jenna Anne Harder

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2020

ABSTRACT

MODELING DECISION PROCESSES IN THE USE OF LETHAL FORCE:
THE ROLE OF RACIAL BIAS IN JUDGING FACES

By

Jenna Anne Harder

To empirically address the question of whether and why police officers are more likely to shoot Black than White suspects, psychologists have developed the First-Person Shooting Task (FPST): a laboratory task in which participants must make shooting decisions based on rapid assessments of whether a Black or White target is holding a gun versus a harmless object. Typically, studies employing the FPST have found that participants' errors and reaction times show a bias toward shooting Black targets over White targets. Evidence for the mechanisms behind this bias is mixed, but several studies point to stereotypic associations between the category "Black" and some indication of threat (e.g. weapon possession). Collectively, this past work is suggestive that racial bias on the FPST is influenced by racial bias in threat perception. I investigated this hypothesis across three studies. Participants rated Black and White faces with regard to how "threatening" the faces appeared, then completed the FPST 3-15 days later. Behavioral and process-level (Drift Diffusion Model) methods were used to determine whether racial bias in a participant's threat ratings explained racial bias in the FPST. Across two stimulus sets, results indicated that although participants displayed process-level racial bias, this was not explained by biased threat perceptions. I consider implications such as the possibility that biased shooting decisions are produced by information-processing mechanisms rather than affective mechanisms.

This dissertation is dedicated to Clementine.
In all these years, you have never spoken an unsupportive word to me.
Thank you for believing in my dreams.

ACKNOWLEDGEMENTS

I am grateful to the Social and Personality Psychology faculty at Michigan State for their support and guidance during my time here. Their influence has valuably improved my abilities to evaluate and produce psychological science. In particular, I would like to thank my committee—Joe Cesario, Debby Kashy, Rich Lucas, and Erik Altmann—for their advice on this work; my dissertation is much stronger because of their contributions. Joe, my advisor, deserves particular recognition for mentoring me throughout graduate school; I am fortunate to have had his influence on the way I think about theory and research design. I'd also particularly like to thank Debby for patiently helping me with countless complicated questions about multilevel modeling over the years. Her instruction and advice have made me a much better statistician. Many thanks as well to David Johnson for his key advice on the diffusion model analyses in this work.

The data for this research were collected by a small army of undergraduates, and I thank Meghan Abrom, David Auer, Kelsey Bees, Erin Brooks, Katie Edwards, Kate Frieden, Joe Healy, Erin Kurth, NaVia Miller, Samuel Motzny, Lauren Scott, and Muhammad Shaharuddin for their time and dedication. This dissertation would not have been possible without them.

Finally, I'm grateful for all the people who have been sources of support and encouragement during my graduate career. Thank you to my parents, Tim and Julie Harder, my brothers, Jim and Mark Harder, and a host of others including (but not limited to): Katie Leahy, Jeewon Oh, Victor Nahuel Felix de Souza Keller, Rosaura Dominguez, Patricia Maina, Mithila Vidyanath, Kiana Gonzalez, Lindsay Weiss, Clementine Theodosia Weiss, and Nathan Wood. Their support has helped me find the energy and confidence to meet the daily challenges of graduate school throughout the last five years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Police shootings of unarmed, innocent Black Americans have drawn widespread attention over the years and incited considerable public controversy. Many people have expressed outrage over what is seen as a pattern evincing widespread racial bias among police, while others argue that protesters are seizing upon isolated incidents in an effort to vilify police officers.

To empirically address the question of whether (and why) police officers are more likely to shoot Black than White citizens, psychologists have developed the First-Person Shooting Task (FPST). This laboratory task requires participants to make rapid judgments to *shoot* or *not shoot* Black and White male targets, using the criterion of whether a given target is holding a gun or a harmless object. Typically, studies employing the FPST have found that participants shoot armed Black targets faster than armed White targets and shoot unarmed Black targets more often than unarmed White targets (Correll, Park, Judd, & Wittenbrink, 2002; for a meta-analysis, see Mekawi & Bresin, 2015). This effect, often referred to as "shooter bias," generalizes across methodological variations. For example, while the classic task shows the entire body of a target holding an object in his hand, Plant, Goplen, and Kunstman (2011) replicated this bias in a task in which pictures of objects simply appear next to Black and White faces. Shooter bias in computerized tasks has been observed in both civilian and police samples (although the findings in police samples are somewhat less consistent, particularly when smaller samples are used; Correll, Park, Judd, Wittenbrink, Sadler, & Keesee, 2007; Sadler, Correll, Park, Judd, 2012; Sim, Correll, & Sadler, 2013).

Evidence for the mechanisms behind shooter bias is indirect and mixed, but several studies point to the role of stereotypes about Black and White Americans. That is, participants may be more likely to shoot Black targets because they associate the category *Blacks* with threatening behavior. Racial bias on the FPST can be amplified by increasing the proportion of stereotype-congruent targets presented in the task, or by arranging for participants to read a “news article” prior to the task about a violent crime supposedly committed by a Black (vs. White) individual (Correll, Park, Judd, & Wittenbrink, 2007; Sim et al., 2013). Moreover, while experience with shooting decisions is usually related to fewer errors and therefore lower bias, an exception to this pattern occurs if officers’ or participants’ experience reinforces stereotypes connecting Blacks with violence or criminality, in which case greater experience is *not* related to lower bias (Sim et al., 2013). For example, Sim et al. (2013) found that patrol officers, who would have received training in making shooting decisions, showed no bias on the FPST. However, a sample of special unit officers from gang and street crime units showed a greater bias than non-police participants; these individuals, too, would have received training in shooting decisions, but also had experiences interacting with gang members of color that may have reinforced stereotypes. Sim and colleagues further found greater bias among civilians who had previously practiced a version of the FPST in which Black targets were more likely than White targets to be armed. Finally, individual differences in shooter bias are related to individual differences in cultural stereotypes: that is, the extent to which an individual believes that “there is a negative stereotype of African Americans as dangerous and aggressive” (Correll et al., 2002). In other words, regardless of whether they claim to personally endorse this stereotype, participants who believe that the stereotype is widespread in society display a stronger bias toward shooting Black targets. This finding, it should be noted, was limited by a small sample and small effect size and was one among many moderators tested, with no correction for multiple comparisons. However, Correll, Urland, and Ito (2006) replicated the finding and found that it was mediated by an event-related potential (ERP) related to perceptions of threat.

The common thread throughout these studies is the association of the category Black with some indication of threat: participants exhibited greater shooter bias if they associated Blacks with violent crime, with the likelihood of having a weapon, or with the characteristics "dangerous" and "aggressive." Collectively, therefore, these findings are suggestive that racial bias on the FPST is influenced by racial bias in threat perception. This interpretation is further supported by evidence that presenting alternative, more direct cues to threat in the FPST stimuli (i.e., depicting targets in neighborhoods that appear "dangerous") eliminates racial bias in participants' responses (Correll, Wittenbrink, Park, Judd, & Goyle, 2011).

Threat Perception

Fully understanding how threat perception can influence racial bias in the FPST requires some understanding of the functions and origin of the human threat perception infrastructure. Researchers have argued that humans have evolved to have cognitive adaptations for threat perception, which include efficiently identifying and rapidly responding to cues of threat. There is an obvious evolutionary advantage to quickly identifying organisms that may attack, as well as to quickly learning what is dangerous without needing large numbers of potentially deadly trials. Animals and humans learn both species-typical defensive reactions (Bolles, 1970) and fear of certain dangers such as snakes and spiders (Öhman & Mineka, 2001) faster than would be expected from typical operant or classical conditioning, indicating the existence of evolved modules that specifically manage ancestrally-recurrent threats.

For humans, one recurrent source of threat has been other humans. That is, throughout evolution, humans have needed to be vigilant against threats from other humans. As such, one should expect humans to be vigilant to signals of impending threat from others, i.e., signals which indicate the intention to harm. The expression of anger is one such signal, as anger may indicate forthcoming interpersonal violence. Accordingly, there is evidence that humans have an evolved module for efficiently detecting signs of anger, which may indicate that evasive or defensive action is necessary in case the angry person behaves violently. A number of studies have found that people can locate a single angry expression in a crowd of faces faster than any other uniquely represented expression (Fox, Lester, Russo, Bowles, Pichler, & Dutton, 2000; Hansen & Hansen, 1988; Horstman & Bauland, 2006; Öhman, Lundqvist, & Esteves, 2001; but see Becker, Anderson, Mortensen, Neufeld, & Neel, 2011, who report a search advantage for happy faces instead). When simultaneously presented with both a neutral and an angry face, people also orient to the angry face first (Cooper & Langton, 2006).

Ingroup/Outgroup Categorization and Ingroup Favoritism

In addition to the tendency to preferentially attend to anger cues, humans have also responded to selective pressures to identify threatening people by evolving a tendency to categorize others as “ingroup” or “outgroup” members. Sorting others into these categories has implications for threat perception.

Humans readily pick up on cues to group membership (Kurzban, Tooby, & Cosmides, 2001), and use these membership judgments to guide their behavior toward other individuals. For example, people evaluate their ingroups more positively than outgroups (Doise, Csepeli, Dann, Gouge, Larsen, & Ostell, 1972), and show favoritism toward their own groups when allocating resources (Tajfel, Billig, Bundy, & Flament, 1971). Ingroup favoritism in resource allocation is even observed in minimal groups—that is, arbitrarily assigned groups with no history or expectation of meaningful interaction (Tajfel et al., 1971). Research in the minimal group paradigm has shed some light into the mechanisms behind ingroup favoritism, finding that this behavior is related not to outgroup hostility so much as ingroup beneficence (Brewer, 1979), and is mediated by participants' expectation that other ingroup members will favor them in fulfillment of a norm of generalized ingroup reciprocity (Yamagishi & Kyonari, 2000).

Ingroup/Outgroup Bias and Race

The adaptive importance of group membership has also led to the evolution of mechanisms that track whether a person is an ingroup or outgroup member, and these mechanisms have implications for race relations. Researchers have hypothesized that humans have evolved to efficiently detect cues to group membership and categorize others accordingly. Since there are a variety of ways in which people might be marked as members of a given social group—including relatively voluntary cues, like attire, or inborn cues, like family resemblance—these cues may take a variety of forms. Because cues to group membership can also change over time and across cultures, an adaptation to detect such cues would need the capacity to quickly learn whatever cues are relevant in the individual's current social environment, regardless of whether they were relevant when the adaptation first evolved.

In present-day society, individuals' social networks are often disproportionately composed of same-race individuals. That is, controlling for population proportions of different races, two individuals are more likely to be friends, neighbors, or even spouses if they are of the same race. For this reason, Kurzban et al. (2001) argue that the cognitive mechanisms that track cues to group membership pick up on this tendency toward within-race affiliation and quickly learn to treat race as a cue to coalition—which, they posit, may explain much about interracial relations in modern society.

Kurzban et al.'s (2001) evidence for this claim builds on previous work using an experimental paradigm known as the "Who Said What?" task, which has been used to support the claim that people "automatically" categorize others by race. In this task participants are instructed to form impressions of a group of targets as those targets have a conversation. After watching the conversation, participants are given a surprise recall task in which they are shown the statements from the conversation and must indicate which target said each statement. Importantly, half the targets are Black and half are White. Studies using this task (Pietraszewski & Schwartz, 2014; Taylor, Fiske, Etcoff, & Ruderman, 1978) have demonstrated that the race of targets influences the kinds of errors participants make in the recall task. A participant who does not remember which target said a given statement has an above-chance likelihood of attributing that statement to a target of the same race as the true speaker. In other words, “within-race errors” are more likely than “between-race errors.” This is taken to indicate that even without conscious intention to do so, participants have encoded the races of the speakers.

Kurzban et al. (2001) used the Who Said What? task to test their hypothesis that people encode race not for its own sake, but in order to make inferences about group membership. If this hypothesis was true, they reasoned, race would be encoded less strongly in a situation where it was clearly irrelevant to group membership. They therefore modified the task so that the observed conversation clearly indicated that targets were divided between two opposing coalitions, with each coalition composed of 50% Black men and 50% White men. This enabled them to assess both categorization by race (comparing within-race errors to between-race errors) and categorization by coalition (comparing within-coalition errors to between-coalition errors). Not only did participants engage in extensive categorization by coalition, but categorization by race was greatly reduced compared to the classic version of the task. That is, coalitional cues led coalitional categorization to largely *replace* racial categorization. These results indicate that “automatic” race categorization is in fact the outcome of a tendency to categorize others by coalition membership and to use race as a coalitional cue. This has the implication that individuals are likely to treat people of other races as outgroup members absent any other information.

Understanding the coalitional connotations of racial categorization is helpful for understanding racial prejudice and discrimination. In the ancestral environment, when most people lived in relatively small, clearly-defined groups, an interaction with another person was more likely to involve a physical altercation if the other person was an outgroup member (controlling for frequency of contact; Ember, 1978). Wariness of outgroup members therefore conferred a survival advantage. Since race is used as a cue to group membership in modern times, any fear or animosity directed toward other racial groups today may be partially a product of a general bias toward treating outgroup members as potentially dangerous.

This may account for at least part of the substantial body of evidence that racial outgroup members (i.e., individuals of a different race than the perceiver) are perceived as more threatening. Much of this research has focused on face perception. It is well-established that people regularly form judgments of threat from others' faces. Higher facial width-to-height ratio is associated with greater perceived aggression (Carré, Morrissey, Mondloch, & McCormick, 2010); and these perceptions are, in fact, relatively accurate, perhaps due to associations with testosterone levels (Carré & McCormick, 2008; Carré, McCormick, & Mondloch, 2009). People can also predict, with some accuracy, whether a convicted sex offender was guilty of violent behavior; they make these judgments by using facial cues including heavy brow and overall facial masculinity (Stillman, Maner, & Baumeister, 2010). Shasteen, Sasson, and Pinkham (2015) also report that when searching a set of neutral-expression faces for particular facial features, participants show a search advantage for structural craniofacial features perceived as threatening.

Given the role of humans' ingroup/outgroup mentality in race perception, and their readiness to use others' facial features to make threat judgments, it is perhaps not surprising that a target's race influences how threatening the target is perceived to appear. The faces of young Black men capture White participants' attention just like evolved threats such as snakes and spiders (Trawalter, Todd, Baird, & Richeson, 2008), unless the faces are depicted with an averted eye-gaze. Similarly, the P200 ERP, which responds to cues to the threat of attack such as angry faces, was also observed to respond to Black (relative to White) faces in a sample of non-Black participants (Ito & Urland, 2003), and this racial bias in ERPs predicts shooter bias in the FPST (Correll et al., 2006).

Similar to this relationship between race and perceived threat, or perhaps because of it, is a relationship between race and perceived anger. Evidence suggests that angry expressions are processed differently on Black faces than on White faces. For example, Hugenberg and Bodenhausen (2003) asked participants to watch faces change from a hostile to a happy expression, or from a happy to a hostile expression, and indicate at which point the hostile expression had ended or started. White participants high in implicit prejudice perceived Black faces as having sooner onset and later offset of the hostile expression relative to White faces. Moreover, under conditions of cognitive load, White participants have better facial recognition for angry Black faces than for angry White faces, despite exhibiting an opposite bias (i.e., better recognition for racial ingroup faces) for neutral faces (Ackerman et al., 2006). White participants high in implicit racial prejudice are also more likely to judge a racially ambiguous target as Black if the face is angry (Hugenberg & Bodenhausen, 2004). Interestingly, Dunham (2011) provides evidence from a minimal-groups experiment that the latter finding may not be specific to racial bias, but instead may reflect a pattern by which angry faces are more likely to be categorized as outgroup members in general, further supporting the idea that racial biases in threat perception may stem from ingroup/outgroup bias.

There is some variation in the robustness of these studies which must be acknowledged. In particular, the studies on attentional bias toward Black faces (Trawalter et al., 2008) and judgments of Black faces as angry (Hugenberg & Bodenhausen, 2003) and of angry faces as Black (Hugenberg & Bodenhausen, 2004) were all characterized by small samples and barely-significant p-values, so the results from these studies should be treated with some caution. However, results from the ERP studies (Correll et al., 2006; Ito & Urland, 2003) and the memory bias study (Ackerman et al., 2006) were better-powered and much more robustly significant (*p*s less than 0.005).

Overall, therefore, this work suggests that categorizing someone as a racial outgroup member has implications for judgments about that person's level of anger and the amount of threat the person poses. These perceptual biases may in turn underlie racial bias in behaviors related to threat responses. For this reason, I hypothesize that racial bias in perceived facial cues to threat—hereafter abbreviated as "threat bias"—explains significant variance in racial "shooter bias" exhibited in the FPST.

Analysis of Shooter Data

The question of how to model the relationship between threat perceptions and shooting behavior is complicated by the fact that past researchers using the FPST have employed a variety of analytic techniques. Participants' error rates are often used as the dependent variable, averaged across trials and compared across target type (Black/White, armed/unarmed). This comparison is often performed using either ANOVAs (Correll et al., 2002; Correll et al., 2011; Miller, Zielaskowski, & Plant, 2012; Plant et al., 2011) or signal detection theory (SDT; Correll et al., 2002; Correll et al., 2011; Kenworthy, Barden, Diamond, & del Carmen, 2011; Plant et al., 2011). Response latencies for correct trials are also often analyzed, typically using regression or ANOVA (Correll et al., 2002; Correll et al., 2011).

These analytic approaches can be divided into approaches that focus on behavioral data (errors and response times) and approaches that focus on cognitive processes. It is important to assess data on shooting behaviors themselves, especially error rates, given that this reflects the life-or-death outcome that the FPST is intended to simulate and study. Nevertheless, it is also important to understand the processes behind shooting decisions.

The most common analytic approach used to examine shooting decision processes is Signal Detection Theory (SDT). SDT (Stanislaw & Todorov, 1999) is used to analyze tasks in which a participant discriminates between two possible stimulus types: one in which some “signal” is present (in this case, trials in which the target holds a gun), and one in which no signal is present (trials in which the target holds some other object). In a given trial, the participant decides whether a stimulus represents a signal or not based on the degree of some internal response, representing the perception that a shoot decision is or is not appropriate. The intensity of this internal response may vary: participants may be more or less certain that a shoot decision is appropriate. However, they must make the binary decision of whether this internal response is sufficiently intense to warrant a shoot decision or not. Therefore, if the intensity of the internal response is above a point referred to as the “criterion,” the participant identifies the stimulus as signal and shoots the target; if not, the participant identifies the stimulus as not a signal and does not shoot the target. The intensity of the participant’s internal response will vary even when no signal is present. This is because various factors can create “noise,” increasing the perception of a signal. Noise can be created by stimulus features, such as the particular angle at which a non-gun is held, or by the participant's internal state, such as subjective sense of threat. However, the intensity of the participant's internal response will also vary when a signal is present. For example, the internal response may be more intense when a black gun appears against a lighter-colored background than when the gun appears against a dimmer background. In other words, noise also exists in signal trials. Thus, internal responses on signal trials vary, forming a "signal-plus-noise distribution," and internal responses on no-signal trials also vary, forming a "noise distribution." It is possible for these distributions to overlap: some non-gun images may produce a more intense internal response than some gun images. The degree of overlap between these distributions represents the first of SDT's two parameters, the "discrimination" parameter, which is related to participants' rate of success at identifying when a signal is present. The second of SDT's parameters is the "criterion" participants set, i.e., the minimum internal response the participant requires to decide that a signal is present. A higher criterion means that the participant is making more conservative decisions, requiring a more intense internal response before deciding that a signal is present. Different participants may set different criteria, and a participant may set different criteria for some trials than for others based on factors such as certain stimulus characteristics—for example, target race. Researchers applying SDT to shooting decisions have found racial biases in the criterion parameter and concluded that participants are setting laxer criteria for making the decision to shoot when targets are Black (e.g. Correll Park, Judd, & Wittenbrink, 2007; Correll, Park, Judd, Wittenbrink, et al., 2007; Kenworthy et al., 2011).

Despite its widespread use, though, SDT is not the best option for modeling the cognitive processes behind shooting decisions. As Pleskac, Cesario, and Johnson (2018) show, SDT does not appropriately characterize the nature of the shoot/don't shoot decision. In addition, SDT cannot answer a number of questions regarding response time data in the FPST. SDT uses only error data and cannot detect or explain cases in which race bias appears in reaction time data but not in error data, as is the case when long response windows are used, or in some police samples (Correll, Park, Judd, Wittenbrink, et al., 2007; Sim et al., 2013).

FPST data can be more fully modeled by the Drift Diffusion Model (DDM; Ratcliff & Rouder, 1998), a model of two-choice decision making which uses error and reaction time data to break the participant's decision into its component processes. The DDM can identify whether or not a manipulation changes each of three parameters of interest, which can all alter responding at the behavioral level in various ways but reflect different cognitive processes involved in the decision and occur at different stages in the decision process. The DDM, therefore, can offer much more detailed information about the role that a given variable (such as target race) plays in the shooting decision. Applications of the DDM to FPST data (e.g. Correll et al., 2015; Pleskac et al., 2018) have yielded consistent conclusions about how race influences the decision process, and posterior predictive checks have indicated that these models show good fit to the data. That is, the models developed in these DDM analyses can be used to simulate new data that closely matches the original data (Pleskac et al., 2018). The DDM describes the decision process with four parameters, which delineate a process in which the participant has a starting inclination toward one or the other choice and accumulates evidence for the choices until the accumulated evidence reaches a threshold corresponding to one of the choices (see Figure 1), at which point that choice is made.

The first parameter is beta, the starting point: the location between the two decision thresholds from which the participant begins to accumulate evidence. If the starting point is closer to the *shoot* threshold than to the *don't shoot* threshold, then the participant is initially inclined toward a *shoot* decision even before evidence accumulation begins. Starting point may represent perceived pay-offs or costs for each decision. In the shooting task, starting points are usually closer to the *shoot* threshold (e.g. Pleskac et al., 2018), presumably because participants are awarded and penalized with points for different kinds of correct and incorrect responses, with a payoff matrix that favors shooting. Likewise, starting points could show racial bias, indicating that the participant or officer was making decisions as though the possibility of incorrectly shooting an unarmed person was more aversive for one target race than the other.

The second parameter is delta, the drift rate—that is, the rate or slope at which the accumulated evidence approaches a decision threshold. If the drift rate for one condition is steeper (i.e., farther from zero in either the positive or negative direction), then participants are accumulating evidence toward a given decision threshold more quickly (relative to conditions with shallower drift rates). This could mean (1) that the stimulus is perceived to contain a greater amount of evidence for that decision, (2) that the stimulus is perceived to contain less evidence for the alternative decision (i.e., the stimulus is less ambiguous), or (3) that the stimulus is easy to process; i.e., that decision-relevant evidence is easily extracted from the visual information. Past shooter studies employing the DDM have found that racial bias affects the drift rate (Correll et al., 2015; Pleskac et al., 2018). Specifically, these studies have found that the drift rate is higher for Black than for White targets, indicating "stronger" evidence for a shoot decision. In other words, the drift rate for armed Black targets ascends more steeply toward the *shoot* threshold than does the drift rate for armed White targets, and the drift rate for unarmed White targets descends more steeply toward the *don't shoot* threshold than does the drift rate for unarmed Black targets. This effect might be best understood as indicating that the target's race is treated as evidence in the decision about the target's object. Since past work has found that racial bias in the FPST is related to drift rate, it was hypothesized in the present work that threat bias would influence shooter bias by affecting participants' drift rates.

The third DDM parameter is alpha, the distance between the two decision thresholds. If the threshold distance is larger, the participant is exhibiting more caution, requiring more evidence before making a decision. Larger threshold distances are associated with more accurate responses and slower response times, so threshold distance reflects the participant's priorities in managing the trade-off between speed and accuracy. A larger threshold distance for Black targets might indicate that participants are concerned about making racially biased decisions and are particularly motivated to make accurate decisions for Black targets.

The final parameter, nondecision time, is not used to assess the influence of manipulated or measured variables, as it typically remains approximately constant across conditions. This parameter is the total amount of time spent on non-decision tasks, such as initial sensation of the stimulus and performance of the motor response. In shooter research, the nondecision time is typically around 300 ms (Correll et al., 2015; Pleskac et al., 2018).

Changes in DDM parameters correspond to certain patterns of behavior-level responding in the FPST. All else equal, steeper drift rates correspond to faster response times and fewer errors. Higher starting points correspond to faster and more accurate responding for *gun* trials but slower, less accurate responding for *non-gun* trials; lower starting points correspond to the opposite pattern. Wider threshold distances, on the other hand, correspond to fewer errors and slower reaction times overall, regardless of object type.

These parameters work together to explain behavioral-level patterns of responding, with changes in any one parameter potentially affecting the influence of the other parameters. For example, if a manipulation widens the threshold by *x* units, then participants in the experimental condition will respond more slowly and accurately because more evidence is needed to reach a threshold from the starting point. However, if the same manipulation *also* raises starting point by $x/2$ units, then the distance between the starting point and the *shoot* threshold will be unchanged, whereas the distance between the starting point and the *don't shoot* threshold will be greater by x units. In this case, then, the manipulation will seem to affect only *don't shoot* responses at the behavioral level. Conversely, if behavior-level responses in one condition are slower and more accurate relative to the comparison condition for *only* gun *or* non-gun trials, then this must indicate a change in threshold offset by a change in starting point. (The same pattern could not be explained by a change in drift rate for trials with that object, because a shallower drift rate would lead to slower but *less* accurate responding.) On the other hand, if a manipulation made responses slower and more accurate in both gun and non-gun trials, but to a greater extent for one object type, this would indicate that both threshold and starting point had been changed, but that the starting point had not changed enough to offset the threshold change for one object type. The DDM can therefore use behavioral data to understand what happens at the process level, identifying the combination of parameter values that best predict the observed data.

The Present Research

The broad goals of the present research are: (1) to assess the extent to which racial bias in the judgments of the threat level conveyed by a target's face are related to racial bias in shooting decisions in the FPST, and (2) to understand the cognitive processes behind how these threat judgments explain variance in shooting decisions. The most basic question to be addressed is the question of whether bias toward perceiving Black faces as more threatening is predictive of a bias toward shooting Black targets. Secondary questions of interest include whether targets whose faces tend to be perceived as more threatening are more likely to be shot, and whether the threat level of targets' faces moderates racial bias in shooting decisions.

Testing the hypothesis that bias in participants' threat perceptions affects drift rates in the DDM may also have useful theoretical implications. If DDM analyses indicate that participants' degree of threat bias is related to the drift rate parameter in the FPST, then it is likely that facial judgments of threat are involved in perceptions of the strength of the evidence for a *shoot* decision. Since the nature of the shooting task means that the drift rate can be conceived of as measuring the perceived evidential strength for identifying the object as a gun or harmless object, this would suggest that efforts to diminish shooting bias might benefit from specifically focusing on object identification processes and the ways in which they might be influenced by threat bias.

The proposed research will: (1) begin by piloting materials relevant to the subsequent studies, (2) address the two questions of interest (i.e., whether and how biased threat perceptions influence shooting), and finally (3) test whether the results obtained in these studies generalize to a different stimulus set to test the generalizability of results to new targets.

STUDY 1

Research Questions

Study 1 served as a pilot study and was used to clarify procedures for the subsequent studies. Study 1 asked participants to rate how “threatening” the faces of FPST targets appeared in order to obtain information about the distribution of threat ratings and whether they are affected by racial bias: that is, whether participants rate Black targets as having more threatening faces than White targets. These data were used to inform four methodological decisions relevant to the subsequent studies, both of which asked participants to make similar threat ratings before completing an FPST.

First, Study 1 assessed the level of variation in threat ratings. The question to be tested by the subsequent studies was whether shooter bias is related to threat ratings, but it would be impossible to answer this question if there were no systematic variation in threat ratings. One possible source of variation in threat ratings is the identity of the target. Study 1 tested whether targets showed consistency in threat ratings (as measured by the intraclass correlation) across participants. If ratings showed some consistency within target, then they could be treated as a property of the target. If not, anger ratings were to be investigated as an alternative dependent variable.

Second, Study 1 investigated the degree to which threat ratings varied across multiple pictures of the same target. The stimulus set for the FPST used in the present study includes each target twice: once holding a gun, and once holding a harmless object. If differences in a target’s pose or facial expression between these two pictures have a strong influence on his perceived level of threat, then perceived threat must be analyzed at the level of the individual stimulus image; if not, it can be analyzed at the level of the target individual, which would simplify the procedure and analyses for the subsequent studies. Study 1 addressed this question.

Third, Study 1 investigated the extent to which individual participants showed varying degrees of threat bias. If threat bias varied significantly across participants, subsequent studies would be able to treat threat bias as an individual difference variable, making it possible to test whether individuals' levels of shooter bias can be predicted from their levels of threat bias.

Finally, Study 1 compared two different survey procedures for measuring perceived threat: one in which participants simply rated how "threatening" each shooter target looks, and one in which these items were intermingled with a variety of distractor questions. If the relationship between Black and White ratings was comparable across the two procedures, the simpler version was to be used in the subsequent studies. If not, the version with the distractor items was to be used.

Method

Participants. A sample of *N* = 259 students from the Michigan State University participant pool participated in return for credit in their psychology courses. Thirty-one of these were excluded for failing one of the questionnaire's three probe items (see "Materials"), leading to a final sample of 228. This sample is adequate (see Appendix C) to detect with more than 96.4% power whether the task version alters the relationship between Black threat ratings and White threat ratings by at least a 0.03 change in standardized slope.

Of the participants in Study 1, 157 identified as White, 29 as Black, 23 as Asian, four as multiracial, and 12 as some other race. Three participants declined to specify race. Thirty-five identified as men, 191 as women, and two as some other gender. Participants ranged in age from 18 to 35, with a mean of 19.13 (SD = 1.56).

Materials. In both conditions of Study 1, participants rated a set of 80 pictures. This set of pictures features 20 Black men and 20 White men previously photographed in our lab (see Figure 2). The men wear identical outfits (gray t-shirt and blue jeans) and are each pictured twice: once holding a gun and once holding a harmless rectangular object, such as a calculator or wallet. Each target holds one of four poses (counterbalanced across race and object type) and appears to stand in one of a variety of neighborhood scenes. When photographed, targets were instructed to adopt a neutral expression. Overall, targets were relatively successful at this, though there is nevertheless some variation in their expressions. For the purposes of Study 1, each of these 80 images were cropped to show only the head and upper shoulders of the target. Participants rated these cropped images one at a time in a survey administered online.

Additionally, in Condition 2 of Study 1, this set of images was expanded to also include head-and-shoulder images of a variety of additional people, including: 5 White men, 10 Latino men, and 5 Asian men. In this condition, it was expected that the variety of races and the uneven numbers of targets per race would disguise the purpose of the survey.

Procedure. In Condition 1, participants completed two blocks of ratings, in counterbalanced order. In one block, participants viewed in random order each of the 80 images, and answered the question, “How threatening does this person look?” In the other block participants viewed the same 80 images, again in random order, and answered the question, “How angry does this person look?” Questions were answered using a sliding scale, with endpoints labeled (for example) “Not threatening at all” and “Very threatening.”

In Condition 2, participants rated targets not only on how “angry” and “threatening” they look, but also on the attributes “happy,” “healthy,” and “embarrassed.” In this condition, different attributes of interest were not presented in different blocks; instead, the 300 attribute/target combinations were simply presented in random order.

At the end of the survey, participants answered three true-or-false probe items: (1) "I understood all the questions I answered." (2) "I answered some or all of the questions randomly or dishonestly, without trying to give an accurate answer." (3) "I read all the instructions in this survey." The second of these probe items, in addition to the response options "True" and "False," had the third option "I don't understand this question." The third probe item had the third response option "I read everything except the consent form at the beginning" in addition to the "True" and "False" options. To be included in the study, participants had to answer "True" to all three questions (with "I read everything except the consent form" considered as practically equivalent to "True").

Results

Descriptive statistics for Study 1 can be found in Table 1.

Note regarding missing data. It was discovered after the study that the incorrect stimulus—specifically, a repeat of another stimulus—had been uploaded for one of the threat-rating items, such that participants rated one stimulus twice and did not rate another stimulus at all. Therefore, ratings are missing for one armed Black target. The additional ratings for the repeated stimulus were deleted. Unfortunately, some of the Study 1 materials were repeated for Study 2, and the error was not discovered until after that study; as a consequence, this error also affects Study 2.

Confirmatory analyses. The first question to be addressed by this study was whether threat ratings would show within-target consistency. A within-subjects ANOVA was conducted with no fixed effects and with random intercepts by target, and the intraclass correlation was calculated to determine whether target was a significant source of variance in threat ratings. There was a significant intraclass correlation among threat scores assigned to the same target (*ICC* = 0.08, $p < .001$). The intraclass correlation is equal to the proportion of variance in threat ratings attributable to the random effect of target and can be conceptualized as the average correlation of threat ratings among all possible pairs of participants. This finding indicates that when multiple participants rated the same stimulus, they tended to assign more similar ratings than would be expected from chance. It was therefore concluded that analyses in the following studies should control for variance in threat scores due to the target.

The second question of interest was whether a given target's threat ratings were typically consistent across the two pictures of that target. That is, because each target provided two images (once armed, and once unarmed), and the face was cropped from both of these images to be rated in Study 1, it was of interest whether two pictures of the same target would be given similar threat ratings. To assess this, for each participant a difference score was calculated for each pair of images, and a multilevel model was developed predicting these difference scores from target (as a categorical factor) controlling for the sum of the image pairs' threat ratings, with intercepts varying randomly by participant. Target was a significant predictor of the difference between two pictures' threat ratings (Table 2, $F(38, 8320.13) = 7.67$, $p < .001$), indicating that some targets had higher threat ratings in one picture than the other, and that this happened to varying extents across targets. Therefore, it was concluded that subsequent studies should include both pictures of each target and that analyses would need to be conducted at the stimulus level rather than the target level—in other words, threat ratings could not be collapsed across the two stimuli in which a given target appeared.

The third question was whether threat bias could be treated as an individual difference variable. To test this, a multilevel model was developed predicting threat rating from target race and target object, allowing intercepts to vary randomly by stimulus and by participant, and allowing slopes for the race of the target individual to vary randomly by participant. Results can be observed in Table 3. The key effect was the variance of slopes for target race at the participant level. These had a variance of 52.75 ($p < .001$), indicating that there was significant variation in the effect of target race across participants. Thus, although participants did not show racial bias in threat perception on average ($b = .03$, $t(180.27) = .032$, $p = .975$), threat bias could be treated as an individual difference variable.

Finally, a linear regression was calculated to test whether participants' threat ratings for Black targets predicted their threat ratings for White targets in the same way across the two versions of the survey (one with distractor questions, and one without). The model predicted participants' mean White target ratings from their mean Black target ratings, survey condition, and the interaction of these two predictors. Results can be observed in Table 4. The effect of Black ratings did not significantly interact with condition ($b = .001$, $\beta = .001$, $t(224) = 0.031$, $p = .975$), indicating that this relationship was consistent across survey versions and that the shorter survey version could be used in Studies 2 and 3.

Exploratory analyses. In addition to these analyses, a number of exploratory analyses were conducted. The relationship between threat and object was examined to determine whether faces cropped from targets holding guns differed systematically in apparent threat from faces cropped from targets holding harmless objects. This could have happened if the individuals photographed to create these stimuli had assumed different facial expressions depending on the objects they held. To test this, a multilevel model was estimated predicting threat from object, race, and their interaction. Random effects included random intercepts by participant and target, slopes for race varying randomly by participant, slopes for object varying randomly by target, and all covariances. Neither object, race, nor their interaction significantly predicted threat (*p*s $\geq$ .113; see Table 5).

In addition, the relationship between anger ratings and threat ratings was explored. Anger and threat ratings were correlated at $r = 0.499$. To further investigate anger ratings, a multilevel linear regression (Table 6) was estimated predicting anger ratings from threat ratings, object, race, and the interaction of threat and race. Random effects included random intercepts and threat slopes by both participant and target, as well as random race slopes by participant, random object slopes by target, and all covariances. Anger was predicted by threat overall ($b = .30$, $p < .001$) as well as by the interaction of threat and race ($b = .04$, $p < .001$). Follow-up analyses indicated that the relationship between threat ratings and anger was stronger ($b = .33$, $p < .001$) for Black targets than for White targets ($b = .26$, $p < .001$), perhaps indicating that perceived anger is interpreted as more threatening on a Black face than on a White face.

STUDY 2

Method

Question. Study 2 assessed whether participants' levels of threat bias predict racial bias in their shooting behavior. It also tested whether threat bias influenced drift rates for shooting decisions, and whether threat bias moderated the effect of race on drift rates. The study was conducted in two parts: an online survey in which participants viewed images of targets and provided threat ratings followed by a laboratory session in which participants completed an FPST featuring the same targets.

Participants. A sample of 732 students from the Michigan State University participant pool participated in Part 1 of the study, and a sample of 310 students participated in Part 2. Of the Part 1 participants, 89 were excluded for failing at least one of the three probe items (e.g. "I read the instructions"—see Study 1 Materials) and four were excluded for indicating that they recognized the face of one of the actors. Of the Part 2 participants, five were excluded for indicating that they had not completed Part 1 as well as three who left the lab early without finishing the shooter task. Data from the two parts were matched by means of a series of survey questions (e.g. what is the second letter in your last name?) that, while containing too little information to be identifying, were likely in combination to produce unique identifiers when pasted together. In a number of cases, however, matches were difficult to make due to participants entering different information each time. If one of the resulting identifiers had no exact match but was only one number or letter away from an unmatched identifier for the other part, the two identifiers were treated as matching pair. Using this method, it was possible to match data across the two parts for a final sample size of $N = 221$.

Participants received credit for their psychology courses for each portion of the study. Of the 221 participants, 153 identified as women, 62 as men, and 2 as a different gender, while 4 did not report gender. Seven did not report race; of the remaining participants, 139 identified as White, 26 as Black, 35 as Asian, 4 as multiracial, and 10 as a different race. Participants ranged in age from 18 to 24, with a mean age of 19.11 (*SD* = 1.18).

The sample size was selected based on results from simulations (see Appendix C for commented simulation R code) indicating that a sample size of 200 would provide 98.4% power and 96.8% power, respectively, to detect a standardized slope of 0.008 for the three-way (threat bias x race x object) interaction terms in the two behavioral analyses of interest described below. The target sample size was therefore 200. Participants were recruited based on rough estimates of the expected attrition rates and of the percentage of Part 2 data that would be usable, e.g. based on matchability to Part 1 data (this estimate was informed by examining the rate of usable data partway through data collection). This resulted in a total of 221 participants. This sample size is also comparable to those of previous studies that have employed DDM analyses of similar complexity (e.g. Correll et al., 2015; Pleskac et al., 2018).

Materials and procedure. Study 2 consisted of two tasks. The first task in Study 2 was the same threat-rating questionnaire used in Study 1, but this time, the task was limited to judgments of threat for each of 80 FPST stimuli. This survey was administered online.

Participants completed the second task in the laboratory, 3-15 days after completing the first task. This task consisted of a shooter task, run in PsychoPy (Peirce et al., 2019) and using the 80 stimulus images rated in Study 1. Participants were instructed that if a target had a gun, that person was dangerous and the correct response was to press the *shoot* key, whereas if a target was holding a harmless object, that person was not dangerous and the correct response was the *don't shoot* key. Participants gained and lost points depending on accuracy. Images disappeared at the end of the response window, and participants who failed to respond before the image disappeared would incur the greatest possible loss of points. Each image appeared once in each of two blocks, for a total of 160 trials; in addition to this, participants completed 8 practice trials at the beginning of the task. The response window for this task was 650 ms.

Many of the actors in the stimulus images were undergraduates who had been photographed in our lab. Partway through data collection, it came to my attention that one of these individuals had passed away, after a participant who knew him contacted me. To minimize potential distress to participants who might recognize him, the images of that individual were replaced with new images of a different individual for the remainder of the study. Eighty-seven participants completed the study with the first version of the stimuli (with the deceased individual) and 134 completed the study with the second version of the stimuli (with a replacement for the deceased individual).

Analyses

Data cleaning. Certain adjustments were made to the data prior to analyses. First, as in Study 1, the incorrect picture had been uploaded for one of the threat-rating items in Part 1, such that participants rated one stimulus twice and did not rate another stimulus at all. Threat perceptions for the unrated target were therefore unknown. Again, the repeated ratings were deleted.

Second, a disadvantage of the survey software used (Qualtrics) was that if a respondent does not move a slider scale from its default start position, the software records this as a nonresponse ("NA"). Threat was rated in Part 1 using a slider scale with values from 0 to 100, with a default starting point of 50. If participants did not move the slider scale for a picture, it was therefore unclear if they intended to respond “50” or not at all. Due to the randomized nature of the survey, “NA” values could also indicate questions that appeared at the end and remained unseen because the participant quit partway through. Visually salient instructions were provided briefly explaining this to participants and asking them to click on the slider scale if they wished to respond with the scale’s midpoint. However, it cannot be guaranteed that all participants read, understood, and recalled all instructions. Therefore, “NA” values for ratings were replaced with “50” only for participants who completed the full survey up to and including the last few pages (demographics, probe etc.). As a result, 210 “NA” values (1.18% of ratings) were replaced with “50.”

For each target, I computed the average of all participants’ threat ratings. For each participant, I computed the difference between (*a*) the participant’s average threat ratings for all Black targets and (*b*) her average threat ratings for all White targets, as well as the sum of these averages ($a + b$).

Behavioral analyses. Descriptive statistics for the threat-rating portion of Study 2 can be viewed in Table 1, and descriptive statistics for the shooting task portion of Study 2 can be viewed in Table 7. Behavioral analyses involving reaction time were estimated using only trials with correct responses, excluding trials on which the participant responded faster than 300 ms or slower than 650 ms (the response window). Response latencies more than 2.5 times the participant’s own standard deviation of response speeds above that participant’s mean were replaced with the value of the participant’s mean plus 2.5 times their standard deviation.

Behavioral analyses, answering the question of *whether* perceived threat influences participants’ shooting behavior, were comprised of four multilevel regression models. The first pair of regressions modeled reaction times and errors, respectively, to assess whether individual participants' levels of threat bias affected the magnitude of shooter bias. The second pair of regressions, which again modeled reaction times and errors, assessed whether the effect of target race on shooting decisions depended on how threatening the target was generally perceived to be (averaged across participants). In each model, random effects (detailed below) were determined based on the most complicated random effects structure that the data would allow without specification errors in a model with no fixed-effect predictors.

The first pair of models treated perceived threat as a property of the participant. These two models were identical in their predictors; however, one was a multilevel linear regression predicting response times, and the other was a multilevel logistic regression predicting errors. Fixed effects in these models consisted of (1) the target's race, (2) the target's object, (3) the difference score for the participant's average Black target ratings minus her average White target ratings, (4) the sum of the participant's average Black target ratings plus her average White target ratings, and (5-11) all interactions that are possible without including both difference scores and sum scores in the same interaction. Intercepts were allowed to vary randomly by participant and by stimulus, and slopes for object were allowed to vary randomly by participant. In the reaction time analysis, slopes for race were also allowed to vary randomly by participant. All covariances were included in both models. The effect of interest in these analyses was the three-way (fixed effect) interaction between the target's race, the target's object, and the participant's Black-White difference in threat ratings—i.e., threat bias. This term would indicate whether a participant's degree of threat bias influenced the impact of race on responses to the targets' objects.

The second pair of models treated perceived threat as a property of the stimulus. These were again a multilevel linear regression of response times and a multilevel logistic regression of errors, again with identical sets of predictors. Fixed effects in these models consisted of (1) the target's race, (2) the target's object, (3) the target's average threat rating across participants, and (4-7) all two- and three-way interactions. Intercepts were allowed to vary randomly by participant and by target, and slopes for object were allowed to vary randomly by participant. In the model of reaction time, slopes for race were also allowed to vary randomly by participant. All covariances were included in both models. The effect of interest in these analyses was the three-way (fixed effect) interaction between threat rating, race, and object. This term would indicate whether "threatening" features (e.g., high facial width-to-height ratio) influence the impact of race on responses to the targets' objects. For example, race may be less influential in determining shooting responses among "baby-faced" individuals, because these individuals may be perceived as nonthreatening regardless of race.

Given the length of the response window, it was expected that bias would emerge in the error analyses but not the reaction time analyses, as this window was likely to limit variation in response time. The error models, therefore, were the analyses of interest. However, reaction time data were also modeled for the sake of consistency with past work.

Some notes on the interpretation of behavioral results. Because only reaction times for correct responses are included in reaction time analyses, the test of the main effect of object is equivalent to a test of whether correct *shoot* responses are faster than correct *don't shoot* responses. Thus, if—for example—threat bias interacted with object, this would mean that participants who perceived Black faces as more threatening differed from participants lower in threat bias in their speed discrepancy between correct *shoot responses* and correct *don't shoot* responses.

A “shooter bias” effect emerges when race (Black / White) and object (gun / non-gun) interact to predict either errors or reaction times, such that the effect of a gun (vs. a non-gun) is different for Black than for White targets. If race and object interact to predict reaction times, this indicates that the discrepancy in speed between correct *shoot* responses and correct *don’t shoot* responses differs between Black and White targets. If race and object interact to predict the odds of an error, the interpretation is affected by the fact that the odds of error when the object is a gun represent the odds of shooting unarmed targets, whereas the odds of error when the object is a non-gun represent the odds of failing to shoot armed targets. Therefore, shooter bias for error data indicates that the odds of shooting unarmed targets and/or the odds of failing to shoot armed targets are different for Black than for White targets. For either reaction time or error data, if this race-by-object interaction is itself qualified by another variable, such as threat bias—in other words, if there is a three-way interaction between race, object, and threat bias—then this indicates that the level of shooter bias depends on the level of threat bias.

Drift Diffusion Model analyses. DDM analyses, answering the question of *how* perceived threat influences participants’ shooting behavior, were specified in a manner consistent with previous use of DDM analyses in shooter research (Correll et al., 2015; Pleskac et al., 2018), using a Bayesian hierarchical specification of the DDM. This model was estimated using a Markov Chain Monte Carlo simulation in JAGS (Plummer, 2003) with the Wiener module (Wabersich & Vandekerckhove, 2014). The analysis collected 300,000 samples using an adaptive phase of 1,500 and a burn-in of 500. Drift rate was predicted by target race, target object, participants’ difference scores for Black and White threat ratings, and scores for the sums of participants’ Black and White threat ratings, as well as all interactions that were possible without interacting difference scores with sum scores. Because Bayesian methods of inference were used, I report the most credible value for each parameter as well as the Highest Density Interval (HDI, reported in brackets), which indicates the spread of the posterior distribution for the parameter.

Posterior predictive checks were conducted for each condition (Black/White by Gun/Non-gun) for the probabilities of each decision (shoot/don't shoot) and the response latency means and distributions. This procedure uses the model to simulate data and compares the simulated data to the original data to indicate whether the model gives a good account of the data in each condition. Results indicated that the simulated data closely tracked observed shooting rates. Simulated response times were also distributed similarly to observed response times, although the model somewhat overestimated response latencies relative to the observed data (with the exception of correct *don't shoot* responses, which the model reproduced more closely) and somewhat underestimated variability in response times for correct responses. Overall, however, the model showed a reasonably good fit to the data.

Results

Behavioral analyses. Recall that one pair of models treated perceived threat as a property of the participant, and a second pair of models treated perceived threat as a property of the stimulus. I discuss each of these separately.

Perceived threat as a property of the participant. Results of the model predicting error (see Table 8) indicated that threat bias did not qualify the race-by-object interaction, and that there was no race bias overall. Interestingly, however, threat bias did qualify the main effects of both race and object. Threat bias interacted with object (e^b = .995, p = .008) such that on non-gun trials, higher threat bias was associated with greater odds of shooting unarmed targets (e^b = 1.007, p = .068), and on gun trials, higher threat bias was associated with lower odds of missing armed targets (e^b =.996, p = .219); in short, participants higher in threat bias tended to shoot more often. Threat bias interacted with race (e^b = 1.002, p = .045) such that for Black targets, higher threat bias was associated with a somewhat greater odds of error (e^b = 1.004, p = .227), whereas for White targets, higher threat bias was unrelated to the odds of error (e^b = 1.000, p = .902). To explore the plausibility that this could arise from participants making more errors for targets whom they found more threatening, a fixed-effects logistic regression model was specified predicting error from race, object, and threat, with threat defined as the specific threat rating assigned to the current trial's target by the responding participant. In this model, higher threat ratings were associated with more errors (e^b = 1.002, p < .001).

In the model predicting reaction time (see Table 9), threat bias did not qualify the race bias effect, and there was no race bias effect overall. However, the sum of a participant's threat ratings did qualify race bias (b = .031, p = .026). Follow-up analyses indicated that participants one standard deviation above the mean in summed threat ratings—i.e., people who tended to assign higher threat ratings overall—showed a race effect for gun trials such that correct *shoot* responses were slower for Black targets (b = .84, p = .803), and a race effect for non-gun trials such that correct *don't shoot* responses were faster for Black targets (b = -3.23, p = .340). Participants one standard deviation below the mean in summed threat ratings showed weak tendencies toward selecting *don't shoot* faster for Black targets on non-gun trials (b = -.79, p = .815) and selecting *shoot* faster for Black targets on gun trials (b = -.56, p = .866). Overall, the interaction seemed to be driven by the emergence of a tendency to select *don't shoot* quickly for unarmed Black targets among participants high in summed threat ratings.

Moreover, threat bias did qualify the effect of object (b = -.190, p = .003), such that on non-gun trials, higher threat bias was associated with slower responses (b = .13, p = .446), and on gun trials, higher threat bias was associated with faster responses (b = -.25, p = .108). Overall, responses were faster to guns than to non-guns (b = -15.873, p < .001).

Perceived threat as a property of the stimulus. In the model predicting error (Table 10, stimulus threat level was not related to the odds of making an error. However, stimulus threat level did interact with object. Overall, object was related to the odds of error (e^b = .869, p = .034) such that participants had a greater odds of error on non-gun trials: that is, participants were biased toward shooting the targets. This interacted with stimulus threat level (e^b = 1.023, p = .017) such that for armed targets, higher stimulus threat level was associated with a greater odds of missing (e^b = 1.025, p = .078), whereas for unarmed targets, higher stimulus threat level was associated with a lower odds of shooting (b = .978, p = .102). In other words, surprisingly, higher stimulus threat level was associated with a greater odds of not shooting on both gun and non-gun trials.

The model predicting reaction time results—again using reaction times from correct responses only, with trials excluded as described above—followed a similar pattern (Table 11). Object was related to reaction time (b = -15.87, p < .001) such that participants' correct *shoot* responses were faster than their correct *don't shoot* responses. This effect was qualified by an interaction with stimulus threat level (b = .77, p = .036) such that for non-gun trials, higher stimulus threat level was associated with faster *don't shoot* responses (b = -.74, p = .143), and for gun trials, higher stimulus threat level was associated with slower *shoot* responses (b = .80, p = .127).

Drift Diffusion Model analyses. DDM results can be seen in Table 12. Overall, participants were biased toward shooting in both drift rate and starting point: Participants had a positive modal drift rate (M = 0.142 [.084, .196]) and a modal starting point above the neutral value of 0.5 (*M* = 0.512 [0.503, 0.522]). Participants' typical nonresponse time was estimated to be about 95% of the individual's minimum reaction time (*M* = 0.949 [.947, .950]). The modal threshold distance was *M* = 1.067 (HDI: [1.048, 1.088]).

There was a main effect of race on drift rate such that overall, drift rates were more positive—i.e., biased toward the *shoot* decision—for Black targets (*M* = .185 [.123, .244]) than for White targets (*M* = 0.097 [.037, .157]). There was a main effect of object on drift rate such that overall, drift rates for gun trials (*M* = 1. 073 [1.015, 1.134]) were steeper than drift rates for non-gun trials (*M* = -.802 [-.859, -.743]). Race and object also interacted to predict drift rate such that for armed targets, drift rates were steeper for Black targets (*M* = 1.158 [1.090, 1.224]) than for White targets (*M* = 0.990 [.921, 1.056]), whereas drift rates for unarmed targets did not differ across race. The racial bias in drift rate was expected in the present work, as it is consistent with past work applying the DDM to shooting decisions.

Threat bias had a main effect on drift rate such that participants who were biased toward perceiving Black faces as more threatening were also prone to interpret evidence as stronger for a *shoot* decision, *b* = .020 [.005, .034] (See Table 12 note for comment on units of DDM slopes). However, the sum of a participant's threat ratings was not related to drift rate, and neither of these individual difference variables interacted with any of the other variables. Thus, the hypothesis that threat bias would interact with racial bias (i.e., the race-by-object interaction) was not supported.

Summary. In sum, higher threat bias was associated with a tendency to shoot more often and with a tendency to make more errors for Black targets in error analyses. Reaction time analyses indicated that participants who were high in perceived threat overall were particularly quick to select *don't shoot* for Black targets. Threat bias did not qualify racial bias overall for either error or reaction time analyses, neither of which found evidence that participants were showing a racial bias. DDM results, however, showed a racial bias in drift rates, although they did not indicate that threat bias affected the degree of this racial bias.

STUDY 3

Method

Question. Study 3 replicated Study 2 with a different stimulus set, to test the generalizability of the previous studies' findings across stimuli.

Participants. A sample of 451 students at Michigan State University participated in Part 1 of the study, and a sample of 321 students participated in Part 2. Of the Part 1 participants, 50 were excluded for failing at least one of the three probe items (e.g. answered False to "I read the instructions"—see Study 1 Materials). Of the Part 2 participants, four were excluded for indicating that they had not completed Part 1. Data from the two parts were matched using the same process used in Study 2. With this method, it was possible to match data across the two parts for a final sample size of $N = 233$.

Participants received credit for their psychology courses for each portion of the study. Of the 233 participants, 160 identified as women, 72 as men, and 1 as a different gender. Three did not report race; of the remaining participants, 156 identified as White, 25 as Black, 31 as Asian, 5 as multiracial, and 13 as a different race. Participants ranged in age from 18 to 24, with a mean age of 18.98 ($SD = 1.18$).

The process of determining the target sample size (200) was the same as Study 1. According to results from the simulations used to make this determination (see Appendix C for commented simulation R code), a sample size of 200 would provide 98.4% power and 96.8% power, respectively, to detect a standardized slope of 0.008 for the three-way (threat bias x race x object) interaction terms in the two behavioral analyses of interest described below.

Materials and procedure. The methods for Study 3 replicated those of Studies 1 and 2, but the stimulus set used in this study was a set of pictures developed by Joshua Correll (Correll et al., 2002). Correll's stimuli include 10 Black and 10 White men, each appearing twice with a gun and twice with a harmless object, for a total of 80 stimuli. This stimulus set is similar to the one used in Studies 1 and 2 in that targets appear in varying poses and with neutral expressions; however, unlike the stimuli used in the studies above, these targets were pictured in their own clothes, and outfit is therefore not controlled across race.

Analyses. Study 3 repeated the analyses conducted in Study 2. It also repeated the analyses of stimulus ratings conducted in Study 1 to test whether this stimulus set has a similar variance structure to the one used in Studies 1 and 2. Additional analytic details are provided in the Results section, particularly when analyses diverge in some way from Studies 1 and 2.

Results

Descriptive statistics for the threat-rating portion of Study 3 can be found in Table 1, and descriptive statistics for the shooting task portion of Study 3 can be found in Table 7.

Variance structure. The first question to be repeated from Study 1 regarding variance structure was whether threat ratings would show within-target consistency. A within-subjects ANOVA was conducted and the intraclass correlation was calculated to determine whether target was a significant source of variance in threat ratings. There was a significant average correlation among threat scores assigned to the same target ($ICC = 0.11$, $p < .001$). As in Study 1, therefore, it was concluded that analyses in the following studies should control for variance in threat scores due to the target.

The second question was whether a given target's threat ratings were typically consistent across different pictures of the target. In this stimulus set, there were four pictures of each person: two armed, and two unarmed. To assess consistency across object, for each participant/target combination a difference score was calculated representing the difference between the sum of the two armed pictures and the sum of the two unarmed pictures. A multilevel model was developed predicting these difference scores from target (as a categorical factor) controlling for the sum of the four images' threat ratings, with intercepts varying randomly by participant. Target was a significant predictor of the difference between armed and unarmed pictures' threat ratings (Table 13, $F(19, 9065.57) = 31.95$, $p < .001$), indicating that some targets had higher threat ratings in armed vs unarmed pictures, and that this happened to varying extents across targets. To assess consistency within object, for each participant/target/object combination, a score was calculated representing the absolute value of the difference between the two pictures. A multilevel model was developed predicting these scores from target (as a categorical factor) controlling for object and the sum of the image pairs' threat ratings, with intercepts varying randomly by participant. Target was a significant predictor of the difference between the image pairs' threat ratings (Table 14, $F(19, 9079.15) = 6.20$, $p < .001$), indicating that targets varied in the consistency of these image pairs' threat ratings. Therefore, as in Study 1, it was concluded that subsequent studies should include all pictures of each target and that analyses would need to be conducted at the stimulus level rather than the target level.

The third question was whether threat bias can be treated as an individual difference variable. To test this, a multilevel model was developed predicting threat rating from target race and target object, allowing intercepts to vary randomly by stimulus and by participant, and allowing slopes for the race of the target individual to vary randomly by participant. Results can be observed in Table 15. The key effect was the variance of slopes for target race at the participant level. These had a variance of 39.75 ($p < .001$), indicating that there was significant variation in the effect of target race across participants. Thus, although participants did not show racial bias in threat perception on average (b = -1.37, t(109.09) = -1.349, p = .180), threat bias could be treated as an individual difference variable. This analysis again replicated the findings of Study 1.

Finally, the exploratory analysis was repeated examining the relationship between threat and object. This analysis was intended to determine whether faces cropped from targets holding guns differed systematically in apparent threat from faces cropped from targets holding harmless objects. In a model identical to that used in Study 1 (see above), neither object, race, nor their interaction significantly predicted threat (ps $\geq$ .291; see Table 16), replicating Study 1.

Behavioral analyses.

Perceived threat as a property of the participant. The first pair of models were identical in their predictors; however, one was a multilevel linear regression predicting response times, and the other was a multilevel logistic regression predicting errors. Fixed effects in these models were identical to those used in Study 2. In both models, intercepts were allowed to vary randomly by participant and by stimulus, and slopes for object were allowed to vary randomly by participant. In the reaction time analysis, slopes for threat bias were also allowed to vary randomly by stimulus. All covariances were included in both models.

Results of the model predicting error (see Table 17) indicated that threat bias did not qualify the race-by-object interaction; in fact, none of the model terms significantly predicted error. To be consistent with Study 2, a fixed-effects logistic regression model was also specified predicting error from race, object, and threat, with threat defined as the specific threat rating assigned to the current trial's target by the responding participant; in this model, however, threat did not predict error. In the model predicting reaction time (see Table 18), threat bias did not qualify the race bias effect, and there was no race bias effect overall. However, responses were faster to guns than to non-guns (b = -17.67, p < .001).

Perceived threat as a property of the stimulus. The second pair of models were again identical in their fixed effects to the models employed in Study 2. In both models, intercepts were allowed to vary randomly by participant and by target, and slopes for object were allowed to vary randomly by participant. All covariances were included in both models.

In the model predicting error (Table 19), stimulus threat level was related to the degree of race bias (i.e., it interacted with the race-by-object interaction; e^b = 1.015, p = .041). Specifically, among White targets, higher stimulus threat level was associated with lower odds of choosing *don't shoot* for armed targets (e^b =.985, p = .185) and was unrelated to errors for unarmed targets (e^b = 1.007, p = .570). Among Black targets, higher stimulus threat level was associated with a lower odds of shooting unarmed targets (e^b = .973, p = .144) and was unrelated to errors for armed targets (e^b = 1.009, p = .486). This effect ran counter to hypotheses. Thus, to explore it further, a fixed-effects logistic regression model was specified testing whether the (non-aggregated) threat rating assigned to a given target by a given participant predicted that participant's shooting behavior for that particular target (in other words, the model had the same fixed-effects structure as the stimulus threat error model, except that "threat" was the specific threat rating for the present participant and targets). This model found the same pattern for White targets; i.e., if a participant found an armed White target more threatening, the participant was more likely to shoot that target. However, the model did not find any object-by-threat interaction for Black targets.

In the model predicting reaction time (Table 20), object was related to reaction time (b = -17.19, $p < .001$) such that participants' *shoot* responses were faster than their *don't shoot* responses. No other terms were significant in this model.

Drift Diffusion Model analyses. Diffusion model analyses were specified in the same way for Study 3 as for Study 2. Analyses collected 201,000 samples with an adaptive phase of 3,000 and a burn-in of 500. Results can be seen in Table 21.

Overall, participants were biased toward shooting in both drift rates and starting point: The modal drift rate was positive (M = 0.090 [.030, .145]), indicating that *shoot* drift rates were steeper than *don't shoot* drift rates. Moreover, the modal starting point was above the neutral value of 0.5 (M = 0.518 [0.508, 0.527]). Participants' typical nonresponse time was estimated to be about 95% of the individual's minimum reaction time (M = 0.948 [.947, .950]). The modal threshold distance was M = 1.072 (HDI: [1.054, 1.092]).

There was a main effect of race on drift rate such that overall, drift rates were more positive—i.e., biased toward the *shoot* decision—for Black targets (M = .156 [.093, .215]) than for White targets (M = .021 [-.041, .081]). There was a main effect of object on drift rate such that overall, drift rates for gun trials (M = 1.155 [1.097, 1.217]) were steeper than drift rates for non-gun trials (M = -.988 [-1.046, -.929]). However, race and object did not interact.

Neither threat bias nor the sum of a participant's threat ratings was related to drift rate. For the most part, neither of these individual difference variables interacted with any of the other variables, and the hypothesis that threat bias would interact with racial bias was not supported. However, threat bias did interact with object, such that threat bias was associated with steeper drift rates for unarmed targets (b = -.014 [-.029, .001]), but was unrelated to drift rate for armed targets (M = -.001 [-.017, .014]), b = .007 [.001, .012].

Posterior predictive checks were again conducted for each condition. Results were very similar to the posterior predictive checks conducted for the Study 2 model. The simulated data closely tracked observed shooting rates and reflected observed response times reasonably well. Exceptions to this were that the model somewhat overestimated response latencies relative to the observed data (with the exception of correct *don't shoot* responses) and somewhat underestimated variability in response times (with the exception of incorrect *don't shoot* responses). Overall, however, the model showed a reasonably good fit to the data.

DISCUSSION

Primary Findings

The primary goal of the present studies was to address whether threat bias is associated with racial bias in the FPST—that is, whether participants' tendency to perceive Black faces as more threatening than White faces was predictive of a tendency to shoot Black targets more than White targets. Across Studies 2 and 3, racial bias did not emerge in behavioral analyses, and this was not qualified by threat bias—defined as an individual's tendency to perceive neutral Black faces as more threatening than neutral White faces. Process-level analyses using the Drift Diffusion Model did find racial bias in drift rates in both Studies 2 and 3, such that drift rates were more positive—i.e., biased toward the *shoot* decision—for Black than for White targets. In Study 2, this was driven by a race difference for armed targets, whereas in Study 3, the race effect in drift rate emerged for armed and unarmed targets equally. However, again, this racial bias did not depend on individual differences in threat bias.

This suggests that racial bias in drift rates is explained by some factor other than perceptions of the target individuals as threatening. Consideration of the structure of the task may be illuminating here. In the FPST, participants' instructions are to press one key if a target is holding a gun and another key if the target is holding any other object. Instructions also indicate that targets holding guns "pose a threat to you" and should therefore be "shot" by pressing the appropriate key. However, the *correctness* of the decision, for which points are awarded and lost, is based on the actual object held by the target, not the participant's subjective judgment of threat. The hypothesis that racial bias in drift rates would be influenced by threat perceptions is premised on the two assumptions that (1) this decision about whether an object is a gun or not is influenced by some assessment of threat from the targets and that (2) this threat assessment is influenced by the target's race. However, it may be that participants experience the task as a visual search and object identification task without a strong threat-assessment component to the decision-making process. If so, racial bias might be better explained by stereotypic associations between the category "Black" and the object "gun," which could speed the object-identification process.

Work with the Weapons Identification Task (Payne, 2001) provides evidence for such a stereotype. In this task, participants must rapidly identify whether or not an object is a gun after being primed with a Black or White face. Participants identify guns faster and more accurately after being primed with a Black face (Payne, 2001; Payne, Lambert, & Jacoby, 2002), suggesting a stereotypic association between Black individuals and guns. On the other hand, work with this task has found that racial bias appears in starting point rather than drift rate (Todd, Johnson, Lassetter, Neel, Simpson, & Cesario, under review)—but this may be an artefact of the structure of the task, in which the face appears before the object (as opposed to the FPST, in which the person and object appear concurrently). Thus, it may be that weapon stereotypes, rather than perceived threat itself, underlie biased shooting decisions in the FPST.

Another possible explanation for the lack of a moderating effect of threat is that the response window in the FPST was too short to allow participants to make judgments about threat. Judgments in the initial threat-rating task were made under no time pressure, but shooting decisions had to be made within a 630 ms window. There are limited data on the precise speed at which different emotions are recognized, and perhaps no data on the speed at which facial cues to threat itself are assessed. However, De Sonneville et al. (2002) present data suggesting that adult response latencies in anger-identification tasks may fall in the range of 600 ms to 950 ms. The present response window, though within that range, is at the low end. In Study Two, threat ratings were somewhat predictive of shooting decisions, as individuals who were higher in threat bias made *shoot* decisions faster and more accurately, and had drift rates biased toward the shoot decision overall. Participants who were higher in overall threat ratings also showed some reaction time bias toward selecting *don't shoot* for unarmed Black targets. However, these effects reflect individual differences and may be related to shooting decisions through some process other than threat perception. That is, individuals who are prone to perceiving threat may also share some other characteristic that produces these shooting effects. Stimulus-level threat ratings might more strongly reflect the effect of immediate perceptions of threat on shooting behavior—but in analyses examining stimulus-level threat ratings, these ratings were typically nonpredictive of error and reaction time and did not often interact with other variables. These null findings are consistent with the possibility that participants were unable to form assessments of threat based on targets' faces before making shooting decisions in the FPST.

Another finding which may be relevant is that race bias consistently emerged in drift rate, but not in behavioral results. There are two possible explanations for this discrepancy. One is that drift rate differences were too small to create detectable behavioral differences. Among the behavioral analyses, race bias was directionally (though nonsignificantly) present in three of the four error analyses and two of the four reaction time analyses. Thus, to some extent the drift rate differences were directionally associated with corresponding behavioral trends—although this was not a perfect association. The other possibility is that some other psychological process(es) not reflected in the present DDM parameter specification may have counteracted the drift rate effect, preventing a behavioral effect. For example, if participants set lower starting points or wider thresholds for Black targets, this might have resulted in equivalent errors and reaction times across Black versus White targets despite the drift rate differences. Since race differences in starting point and threshold were not modeled here due to computational limitations, the present analyses cannot assess this possibility.

Stimulus-Level Threat

A second and more exploratory question was whether the average threat level assigned to a stimulus across participants—that is, a stimulus-level measure of threat—would be associated with racial bias in these decisions, or with the decisions themselves. In Study 2, there was no such relationship; however, in Study 3, higher stimulus threat level among Black targets was associated with fewer *shoot* responses for unarmed targets, and higher stimulus threat level among White targets was associated with more *shoot* responses for armed targets. In Study 2, higher stimulus threat level was associated with a tendency toward the *don't shoot* decision in both errors and response times. This was an unexpected pattern of effects. I would have predicted that higher perceived threat level of a stimulus would be associated with an increased tendency toward *shoot* decisions for both races, with perhaps a stronger association for Black targets. Instead, higher threat level was associated with *don't shoot* decisions for all targets in one study and for Black targets in the other. One potential explanation, discussed in more detail above, is the response window during which participants had to respond. If threat judgments are formed too slowly to influence responses faster than 650 ms, then the observed effects of stimulus threat level and participant threat bias may not truly represent what those variables were intended to measure. Another possible factor is that stimulus threat level was calculated as a stimulus-level variable based on threat ratings aggregated across participants, but there was some missing data in threat ratings. That is, not all participants rated all targets (because sometimes participants did not finish the survey). It is possible that some pattern of missing data in Study Three was responsible for this counterintuitive effect. For example, some targets may have been rated as more threatening by responding participants, while a subset of participants who thought they were less threatening did not respond to the threat questions for those targets yet still contributed *don't shoot* responses toward those targets in the FPST. This is a somewhat speculative explanation, but is supported by the fact that in follow-up analyses, the counterintuitive effects of stimulus threat level for Black targets did not emerge when stimulus threat level was not aggregated across participants.

Additional Findings

Some interesting lower-order effects also emerged. In Study 2, participants who were higher in threat bias made more errors for Black than for White targets. A straightforward explanation is that participants who perceived Black targets as particularly threatening may have responded to the sight of a Black target with some negative emotional state that impeded processing. In other words, people may make more errors for targets they perceive as more threatening. Follow-up analyses indicated that the present data supported this claim; i.e., de-aggregated threat ratings for a target predicted the raters' odds of error for that target in Study 2. Interestingly, however, in Study 3, individuals' threat ratings for particular targets did not predict their odds of error on trials presenting those targets—and neither did threat bias predict differences in the odds of error across target race. The main methodological difference between the studies was that they used different stimulus sets; Study 2 employed stimuli that I had previously created at Michigan State University, whereas Study 3 employed stimuli created by researchers at the University of Chicago and the University of Colorado at Boulder (Correll et al., 2002).

This could indicate some difference between the stimulus sets evaluated in the two studies. The reader may wonder whether the Study 3 stimuli were easy to make shooting decisions about, or hard to make threat judgments about, leading to restriction of range in either errors or threat ratings and hiding an association in Study 3. However, examination of the data reveals that the error rate in Study 2 (29.0%) was similar to that in Study 3 (26.4%) and that the standard deviations of threat ratings were also similar between Study 2 (23.4) and Study 3 (24.2). Another possibility, if perceived threat is not truly related to the odds of error, is that the stimuli in Study 2 may have contained a subset of images with some quality that increased the odds of error while also increasing perceived threat (for example, perhaps blurry images might have this effect). On the other hand, if perceptions of threat do increase odds of error, it may be that Study 3 has a subset of stimuli that counter this pattern: e.g., that have some quality that makes them more error-prone but less threatening (for example, images in which targets appear small relative to the background scenes). There is little research in the published literature that can shed light on whether the threat effect in Study Two is valid or spurious, so it is unclear at this point what to make of the differences between the stimulus sets. However, it does raise interesting questions for further research regarding the effects of perceived threat on decision-making processes.

Another interesting finding that emerged in Study 2 was that participants who were higher in threat bias also made the decision to shoot faster and more often and had more positive drift rates (indicating a bias toward shooting). These findings did not replicate in Study 3, so it is possible that these represent spurious effects. However, if they do accurately reflect reality, they may be a sign that some personality trait underlies both a tendency to perceive Black individuals as more threatening than White individuals as well as a tendency toward making aggressive choices or perceiving objects as weapons.

In Study 2, participants who were higher in overall threat ratings also showed some reaction time bias toward selecting *don't shoot* for unarmed Black targets. Again, this effect was not replicated in Study 3, nor was it reflected in any process-level (DDM) effects. Taken at face value, the finding seems to suggest that there is some individual difference associated both with perceiving threat and with correctly failing to shoot Black targets. It is possible that people who are prone to perceiving threat easily may have been particularly on their guard against shooting unarmed Black targets in an effort not to appear racially biased. There is some literature suggesting that participants who show an indirectly measured motivation to respond in a non-prejudiced way may be able to exert some control over their responses in this task (Glaser & Knowles, 2008; Park & Glaser, 2011). However, whether this motivation is related to the tendency to perceive others as threatening is unknown. Thus, again, it is unclear whether this finding represents a true effect.

A final effect worth mentioning is that in the exploratory analyses from Study 1, ratings of anger showed a moderate-to-strong correlation with ratings of threat. This may suggest that the two measures tap into overlapping constructs—implying that past research on racial bias in perceptions of anger could instead be understood as bias in perceived facial cues to threat. There is some evidence that perceptions of anger are related to perceptions of threat: Shasteen et al. (2015), in their study demonstrating that people have a search advantage for facial features perceived as threatening, found that this advantage was exaggerated in people who also perceived faces with these features to look angrier. This could occur because perceived anger directly increases perceived threat, or because attributing anger to faces that are perceived as threatening for other reasons facilitates the processing effects of perceived threat. Participants may also be conflating perceptions of anger with perceptions of threat because they know that anger can be associated with violent or threatening behavior. Moreover, some facial features may contribute to assessments of both anger and threat; for example, lowered eyebrows are considered characteristic of angry expressions (Kohler et al., 2004) and are also perceived as threatening (Lundqvist, Esteves, & Ohman, 1999).

However, it is more difficult to explain why this relationship between anger and threat was stronger for Black than for White targets. This could indicate that anger is perceived as particularly threatening on Black faces, and/or that less angry expressions are perceived as particularly harmless on Black faces. It could also indicate that threatening facial features (e.g. high width-to-height ratio) on a Black face lead the face to be perceived as angrier, relative to White faces with the same features. It is even possible that facial expression was genuinely correlated with some threatening facial features in our stimulus set, even though actors were instructed to maintain neutral expressions, and that this relationship was stronger for Black faces. A systematic investigation of the stimuli would be necessary to test that hypothesis. The race difference observed here is intriguing, but more research will be needed to understand it. That said, however, it should be noted that the relationship described here was observed based on 20 Black and 20 White faces—a relatively small sample—and is therefore of uncertain reliability. Since perceived anger was not measured in Study 3, it is unknown whether this effect would replicate with that study's stimuli.

Validity Considerations

Several possible concerns could be raised surrounding topics discussed in the present work. These include the robustness of the past findings on which the present studies are based, the validity of the task used here to measure perceived threat, and the interpretation of the drift rate parameter in the DDM. I consider each of these in turn.

Robustness of shooter bias literature. Recent years have seen growing concern with the replicability of social psychological studies (e.g. OSC, 2015). Given these concerns, it seems advisable to carefully examine indicators of empirical soundness when citing past literature. I therefore take some time here to consider the robustness of certain shooter literature effects relevant to this dissertation.

There is fairly widespread support for the finding that civilian participants display shooter bias in the traditional FPST. The claim has been backed up by a meta-analysis (Mekawi & Bresin, 2015), and while there is great variation in sample sizes within the shooter literature, a number of studies have drawn conclusions from reasonably large samples (e.g., N > 100; Correll, Park, Judd, & Wittenbrink, 2007; Hunsinger, 2011; Mekawi, Bresin, & Hunter, 2016; Musolino, 2012; Park & Kim, 2015; Pleskac et al., 2018; Sadler, Correll, Park, & Judd, 2012; Sim, Correll, & Sadler, 2013; Snowden, 2017). However, the robustness of this finding to different analyses is somewhat questionable. Johnson, Cesario, and Pleskac (2018) report that analyses still revealed a shooter bias effect when employing a rigorous random effects structure in a multilevel modeling framework. In contrast, though, Harder (in press) subjected data from 19 shooter studies to various random effects structures in a multiverse analysis and found that studies varied in their robustness to more conservative analysis techniques. Harder concluded that most shooter studies were underpowered for the most rigorous analytic options as a consequence of using too few unique target individuals. Thus, while shooter bias effects are frequently detected, some inappropriate methods may be common in this literature.

The moderators and mediators of shooter bias, unfortunately, are also somewhat uncertain, because many studies addressing these questions remain unreplicated. A few such studies have been subject to replication attempts; for example, two studies (Correll, Park, Judd, & Wittenbrink, 2007; Sim et al., 2013) have found that shooter bias is greater after reading a news article that reinforces stereotypic associations between Blacks and violent crime. One of these studies (Correll, Park, Judd, & Wittenbrink, 2007) found this effect to be highly significant with a moderately large sample; the other (Sim et al., 2013) replicated it in a smaller sample with more modestly significant results (p = .032) and found that it appeared in undergraduates with no prior experience with the task, but not in police participants or in experienced undergraduates. However, this is one of the few manipulations that have been repeated across publications in the shooter literature to test explanations for shooter bias. A number of other studies have also examined a role for stereotypes or have tested other explanations, but these use a variety of measures and manipulations (e.g. Correll et al., 2006; Kenworthy et al., 2011; Ma & Correll, 2011; Ma, Correll, Wittenbrink, Bar-Anan, Sriram, & Nosek, 2013; Taylor, 2011).

Although the basic shooter bias effect has emerged in many studies of undergraduates or other civilians, it is more difficult to discern whether the effect exists for police officers, as only a handful of studies have examined police populations. Of studies that tested for both reaction time effects and error effects, at least one study has found an effect in reaction times but not in errors (Correll, Park, Judd, Wittenbrink, et al., 2007, Study 1) and at least one other has found an effect in errors but not reaction times (Taylor, 2011, Study 1). Others have found neither error effects nor reaction time effects (Correll, Park, Judd Wittenbrink, et al., 2007, Study 2; Sim et al., 2013, Study 1; Taylor, 2011, Study 2), or both error effects and reaction time effects (Ma et al., 2013). Sadler et al. (2012) only report testing for bias in reaction times, but do find a bias toward responding more quickly to armed Black targets and responding more slowly to unarmed Black targets. Sim et al. (2013, Study 2b) only report testing for bias in errors in their sample of 22 Special Unit officers, but do find an error bias toward shooting Black targets. In short, evidence for shooter bias among police is quite inconsistent.

The question of how police compare to civilians in shooter bias is of particular interest. A few studies have directly compared a police sample to a civilian sample. However, results from these comparisons are again mixed. Correll and colleagues (Correll, Park, Judd, Wittenbrink, et al., 2007) found in one study that their civilian sample showed bias in both errors and reaction times while officers showed it only in reaction times, and in another study that the civilian sample showed bias in errors while the police sample did not show bias. Taylor (2011) found in one study that police and civilians alike showed bias in errors but not in reaction times, and in another study that neither police nor civilians showed bias in either errors or reaction times. Unfortunately, the relatively few studies comparing police to civilian samples makes it difficult to draw firm conclusions about any difference between these populations.

Another broad issue with answering this question is that sample sizes vary widely in studies of police and are often smaller than samples in studies of civilians, potentially limiting the reliability of conclusions about this population. Particular attention should therefore be given to the most highly powered studies of police. Two of the most highly powered studies report effects specific to reaction time data (Correll, Park, Judd, Wittenbrink, et al., 2007, Study 1; Sadler et al., 2012). However, these studies also used relatively long response windows, which tend to be associated with reaction time effects rather than error effects (Harder, in press). Another large-sampled study (Ma et al., 2013) used a more moderate response window and found both error and reaction time effects, although this study recruited new police recruits rather than experienced officers. Thus, while there is a certain amount of evidence that highly powered studies of police do reveal shooter bias, it is unclear whether police are more likely to show this bias in errors or reaction times. Overall, studies on shooter bias in police samples are again too few, and their findings too mixed, to suggest any clear conclusions.

A final question is whether the findings reported with the basic task extend to similar tasks. A few studies have shown “shooter bias” effects for one similar task (Plant et al., 2011; Plant, Peruche, & Butz, 2005) that eliminates the neighborhood scenes employed in the FPST and, instead of portraying targets’ full bodies holding objects, displays Black and White faces with objects (guns or non-guns) next to them. These studies typically report reasonably large sample sizes and highly significant results, but shooter bias in some of these studies was limited to the first half of trials, suggesting that its magnitude is diminished with continued practice with the task. Nonetheless, one study employing the same task even found that these effects generalized to police (Plant & Peruche, 2005). However, this study recruited a rather small sample (N = 48), and while effects were highly significant for the first 80 trials of a 160-trial task, they again did not emerge in the second half of trials. Thus, this study may represent weak evidence for the robustness of the effect in police samples. The task may therefore most reliably elicit shooter bias among civilians.

However, James and colleagues (James, Klinger, & Vila, 2014; James, Vila, & Daratha, 2013) tested the shooter bias effect with a more naturalistic task and did not replicate the original shooter bias effect. These studies employed small samples and may have been relatively underpowered. Moreover, the data were originally collected to address questions unrelated to race, and uneven proportions of different races (Black, White, and Hispanic) appeared across their shooting simulation scenarios: further lowering power, as each participant would have responded to relatively smaller numbers of Black targets. Nevertheless, they report results for most of their experiments indicating a bias toward shooting White targets more than Black targets—that is, in the opposite direction of the typical effect. These results were highly significant in two of the three experiments in their first paper (James et al., 2013) and were at the

$p < .05$ level in the one study of their replication paper (James et al., 2014). One caveat regarding this work is that there is no indication that scenario or object were counterbalanced across race. In a more traditional shooter task, each target appears both holding a gun and holding a harmless object across multiple trials, to minimize the influence of individual targets on results. Particularly given the small number of trials (participants completed between 10 and 27 trials depending on the study), the fact that the relationships between race and scenario or between race and object were not controlled is a concern for interpreting these findings.

In conclusion, basic shooter bias effect, with the traditional task and with civilian samples, seems to be the most clearly reliable finding in this literature. Evidence is suggestive that the effect may also extend to police samples. However, it is unclear whether shooter bias appears in tasks that differ substantially from the original FPST paradigm.

Measuring threat perception. A possible limitation of the present work is that threat perception may not have been measured in a valid way. The present measure of threat, though high in face validity, has not been validated in any past work. Past studies assessing racial bias in threat perception have used a variety of tasks, and have often focused on perceived anger instead of measuring perceived threat directly. For example, one study asked participants to identify the new emotion in a facial expression changing slowly from happy to hostile, or from hostile to happy, and measured how long participants considered the expression to be “hostile” for Black and White faces (Hugenberg & Bodenhausen, 2003). Another study asked participants to identify the race of a racially ambiguous face displaying an angry or neutral expression (Hugenberg & Bodenhausen, 2004). Other studies have used psychophysiological measures to detect threat responses to neutral Black and White faces (Correll et al., 2006; Ito & Urland, 2003). The present task, however, simply asked participants to rate how “threatening” each face appeared for many neutral Black and White faces. If perceived threat was not truly measured by the present threat-rating task, this provides an alternative explanation for why variables derived from these ratings did not predict shooting behavior.

There are a few reasons to believe that the threat-rating task in the present studies may have been a valid measure. Threat ratings correlated with ratings of perceived anger in Study One. Moreover, although the task did not resemble any tasks used in past studies of racial bias, it was similar to the method used in Carré et al.'s (2010) study of the correlates of threat perception. Carré et al. had participants rate neutral faces in response to the question "How aggressive would this person be if provoked?" They did find that participants' aggression ratings correlated with targets' facial width-to-height ratios and a behavioral measure of targets' actual aggression, which is suggestive that the task may be a valid measure. However, their task did differ from the present studies' task in that participants made ratings about "aggression if provoked" rather than "threat," which may be a less concrete judgment. Moreover, participants did not, on average, perceive the Black targets' faces as more threatening than the White targets' faces, inconsistent with past work demonstrating racial bias in perceived threat.

To have a clear idea of whether the present task truly measured threat perception, an attempt should be made to validate the specific task used in these studies. A reasonable strategy for doing so would be to evaluate whether individual differences in threat bias on this task correlate with measures of racial bias in perceived threat (or anger) used in past studies, such as the task used in Hugenberg and Bodenhausen's (2004) study on the racial categorization of angry faces. It might be particularly useful to test whether ratings on this task were predictive of neurological responses, such as the event-related potentials studied by Ito and Urland (2003)—and whether racial bias in threat ratings was predictive of racial bias in those neurological responses. If future research applies such methods to validate the present threat-rating task, the present results can be interpreted with greater confidence.

Interpretation of drift rate effects. Another question that may affect the validity of the interpretations presented here is how, exactly, the drift rate parameter of the DDM should be understood. Traditionally, the drift rate is considered to indicate the perceived strength of the evidence for a given decision, based on the tenet that the parameter represents the rate of accumulating evidence for a given decision over time. However, a general limitation of "process models," including the DDM, is that numerical parameters are interpreted as representing specific psychological processes that cannot be directly observed or measured, with the consequence that these interpretations are not directly testable. It is therefore worthwhile to consider what evidence exists to support this interpretation.

Some work has been done to address this question. Ratcliff and McKoon (2008) present several experiments testing predictions about which DDM parameters would respond to certain manipulations. One experiment varied the strength of the visual evidence for each decision and found that, consistent with the conventional interpretation of drift rate, this manipulation caused a change in participants' drift rates such that drift rates were steeper for stronger evidence, but did not affect any other parameters. Other manipulations that bore no implications for evidential strength largely did not affect drift rate, with the exception that increasing base rates (75%/25%) of a particular stimulus type led to slightly higher drift rates for that type (in addition to a more substantive impact on starting point). Analogous conclusions were drawn by Voss and colleagues (Voss, Rothermund, & Voss, 2004), who had participants judge which color was more common in a two-color array and varied the percentage of the more common color as a manipulation of evidentiary strength. Participants' drift rates were steeper when the difference between the two colors’ frequencies was greater, but no other parameters were affected by this manipulation. Other parameters tested in these studies did not affect drift rate, with the exception of a manipulation inhibiting the speed with which participants could physically respond, which affected drift rates for blue (but not orange) stimuli, among other parameters. Similarly, Ratcliff and Rouder (1998) demonstrated across three studies that drift rates were directly related to the probabilities that a stimulus belonged to one or the other category, consistent with a strength-of-evidence interpretation (they do not report whether other parameters were affected by this manipulation). Overall, these studies provide reasonably consistent evidence that drift rate corresponds to some perception of the strength of the evidence for a decision.

However, each of these studies asked participants to differentiate between two types of stimuli that varied in some perceptual quality, but not in psychological significance—for example, dots moving toward either the left or the right of the screen (Ratcliff & McKoon, 2008). The shooter task, on the other hand, asks participants to differentiate between guns and non-guns, which may be experienced differently given that guns may be more salient or affectively significant than harmless objects. Pleskac and colleagues (2018) varied the discriminability of objects in a shooter task and found more mixed evidence for the classic interpretation of the drift rate. Specifically, this study presented some shooter task objects as blurry—i.e., directly manipulating the strength of the visual evidence for the *shoot* decision. As in other studies, results indicated that the evidential strength manipulation (in this case, blurriness) affected the drift rate parameter, but did not alter any of the other parameters. The finer details of this pattern, however, are inconsistently supportive of the strength-of-evidence interpretation: Blurriness was associated with shallower drift rates for gun trials but with steeper drift rates for non-gun trials. The traditional interpretation of the drift rate is consistent with the former effect, i.e., lower drift rates (indicating less evidence for a *shoot* decision) for blurry guns. The latter effect, however, is harder to interpret according to this understanding of the drift rate, as this interpretation would mean that blurriness increased the evidence that an object was a non-gun.

The mixed effects observed by Pleskac et al. (2018) suggest that some nuances of the current understanding of the drift rate may be inaccurate. When the drift rate is applied to shooting decisions, the standard definition of the drift rate implies that participants are accumulating both any evidence that the object is a gun *and* any evidence that the object is some other object. These two sets of evidence are continually pitted against one another as information is gathered until the net evidence for one option reaches the threshold for that option. By this understanding, blurriness should not increase the evidence that the target is holding some non-gun object. Instead, Pleskac et al.'s (2018) findings may be more consistent with a decision process in which participants are making yes/no decisions about whether the object is a gun: i.e., instead of accumulating positive evidence that the object is some specific non-gun item, they may accumulate evidence that the object is a gun and pit this against the absence of such evidence—e.g., if a particular gun-like shape is sought and not found, this may be treated as evidence for the *don't shoot* decision. Such a process would be consistent with the pattern of results reported by Pleskac and colleagues (2018), as a participant seeking a clear gun-like shape might reject a blurry non-gun faster than a clear non-gun.

This interpretation, however, is speculative, and it should be noted that a replication of Pleskac et al.'s (2018) finding would be an important first step in assessing this possibility. It is an interesting issue worthy of further study—however, at least for purposes of interpreting the present results, even this implication of Pleskac et al.'s (2018) findings still supports the broad conclusion that drift rate represents some aspect of assessing the evidence for a decision. I would therefore conclude that the preponderance of evidence supports the interpretations of drift rate results presented here, but would caution readers that some nuances of these interpretations may still be unknown.

Conclusion, Summary, and Future Directions

Many researchers (e.g., Correll et al., 2015; Johnson et al., 2018; Pleskac et al., 2018) have made claims about the role of threat in shooting decisions; however, this was the first research to provide a direct test of that relationship. The literature reviewed above suggested that shooting bias was related to holding stereotypic associations between Blacks and weapons, violent crimes, or labels such as “aggressive” and “dangerous.” This produced the present hypothesis that perceptions of Black individuals as *threatening* would explain shooter bias. On the whole, it appears from the present results that racially biased threat perceptions do not qualify racial bias in shooting decisions; that is, the hypothesis was not supported.

Thus, the question remains open as to what cognitive process underlies the relationship between these stereotypes and shooting behavior. As discussed above, it may be that the relationship is cognitive rather than affective, and is based not on perceived threat so much as on associations between Blacks and weapons. Future research might fruitfully explore this by pitting threat perception against weapons stereotypes as explanations for shooting bias.

The possibility should also be acknowledged that the FPST is too artificial a task to capture the process by which racially biased shooting decisions might play out in the real world. An officer encountering a Black suspect in the course of police work may experience a state of threat that influences subsequent behavior, but this does not guarantee that the same officer encountering a Black target in the FPST will experience this state of threat. This question, of the validity of the FPST itself, is an important one for researchers to consider, though finding a satisfactory answer will be difficult given that real-world shooting decision scenarios are rare and occur under widely varying circumstances. Moreover, future research should examine the timeline across which threat perception unfolds. Too-short response windows could have prevented the detection of a relationship between threat bias and shooter bias, leading FPST results to inaccurately represent shooting decisions in real-world scenarios that unfold over multiple seconds. A better understanding of the time it takes to process threat will shed light on this.

Some of the secondary findings from the present studies suggest additional avenues for future research, as some of the relationships that appeared in the present studies could be better understood with further research on certain topics. First, it may be helpful to clarify the relationship between perceptions of threat and accuracy in decision-making processes such as those involved in the FPST. Better understanding this relationship will shed light on why participants who were prone to perceiving Black targets as more threatening in Study 2 were also prone to making more errors for Black targets. Second, further replication attempts with these and other stimulus sets will assist in determining the reliability and boundary conditions of effects that appeared in only one study. Third, systematically measuring qualities of the stimuli, such as blurriness or the size of targets relative to the frame, would allow for testing whether any of these qualities are associated with particular patterns of responding. This could also prove fruitful for understanding why different results appear with different stimulus sets.

In conclusion, the present research suggests that—despite claims to the contrary in the literature—racially biased perceptions of threat are not a valid explanation for individual differences in shooting behavior on the FPST. The three studies also identified a number of unanticipated secondary findings worthy of replication attempts and further study. Most importantly, however, this research has indicated that shooting decision researchers should reconsider the assumption that decisions on the FPST are a reflection of threat perception. It is time to explore other processes that may underlie shooting bias.

APPENDICES

APPENDIX A: TABLES

Table 1

Descriptive Statistics from Threat-Rating Data for Studies 1, 2, and 3

	Threat	Stimulus Threat	Threat Bias	Threat Sum
Study 1				
Overall	36.19 (27.21)	36.21 (8.02)	.05 (15.15)	72.63 (35.14)
Black	36.13 (26.76)	36.15 (7.30)		
White	36.25 (27.63)	36.27 (8.76)		
Study 2				
Overall	38.30 (23.36)	38.08 (6.83)	-.71 (12.25)	76.67 (31.14)
Black	37.94 (22.58)	37.95 (5.94)		
White	38.65 (24.09)	38.20 (7.64)		
Study 3				
Overall	42.84 (24.29)	42.84 (8.28)	-2.73 (13.16)	85.69 (28.69)
Black	41.48 (23.20)	41.48 (7.25)		
White	44.21 (25.27)	44.21 (9.09)		

Note. Values indicate means, with standard deviations presented in parentheses. Threat: raw threat rating. Stimulus Threat: Mean threat rating assigned to a given stimulus. Threat Bias: Difference between a participant's mean threat rating for Black targets and mean threat rating for White targets (positive numbers indicate bias toward perceiving Black targets as more threatening. Threat Sum: Sum of a participant's mean threat rating for Black targets and mean threat rating for White targets.

Table 2

Multilevel Linear Regression Model Predicting Differences in Participants' Threat Ratings between the Two Pictures of Each Target in Study 1

	Fixed Effects			
	F	*df 1*	*df 2*	*p*
Intercept	.81	1	476.73	.370
Target ID	7.67	38	8320.13	<.001
Sum of Threat Ratings	.01	1	777.85	.927

		Random Effects			
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		374.90	5.815	64.47	<.001
Participant	Intercept	.28	.966	.28	.776

Note. N_{sample} = 228. N_{target} = 39. $N_{observations}$ = 8574.

Table 3

Multilevel Linear Regression Predicting Stimulus-Level Threat Rating in Study 1

	Fixed Effects					
	b	*β*	*SE*	*t*	*df*	*p*
Intercept	-.05	-.002	1.47	-.03	275.50	.974
Race	.03	.001	.96	.03	180.27	.975
Object	-.22	-.008	.91	-.25	76.31	.806

	Random Effects				
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		317.52	3.44	92.37	<.001
Participant	Intercept	307.18	29.24	10.51	<.001
	Covariance	11.51	8.86	1.30	.194
	Race	52.75	5.35	9.86	<.001
Stimulus	Intercept	63.71	10.55	6.04	<.001

Note. N_{sample} = 228. $N_{stimuli}$ = 79. $N_{observations}$ = 17,530. Race refers to race of the target. Threat was mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 4

Linear Regression Predicting Participants' Threat Ratings for White Targets from their Threat Ratings for Black Targets and Survey Version in Study 1

	b	*β*	*SE*	*t*	*df*	*p*
Intercept	-.04		.90	-.05	224	.963
Black ratings	.64	.687	.05	14.18	224	<.001
Condition	-.97	-.052	.90	-1.08	224	.281
Black*Condition	<.01	.001	.05	.03	224	.975

Note. N = 228. *F*(3,224) = 67.78. Condition: Survey version, effects coded (Short version = 1; Long version = -1). Mean White and Black ratings were mean-centered.

Table 5

Multilevel Linear Regression Model Predicting Stimulus Threat Ratings in Study 1

Fixed Effects						
	b	*β*	*SE*	*t*	*df*	*p*
Intercept	36.41	.008	1.70	21.36	115.84	<.001
Object	-.17	-.006	.29	-.58	37.31	.564
Race	-.08	.003	1.34	-.06	49.89	.953
Object*Race	.46	.017	.29	1.62	37.31	.113

Random Effects					
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		319.40	3.46	92.37	<.001
Participant	Intercept	304.25	28.95	10.51	<.001
	Covariance	11.74	8.83	1.33	.183
	Race	52.74	5.35	9.85	<.001
Target	Intercept	62.01	14.40	4.31	<.001
	Covariance	-2.03	2.33	-.87	.383
	Object	2.47	.74	3.34	<.001

Note. N_{sample} = 228. $N_{stimuli}$ = 79. $N_{observations}$ = 17,530. Race and Object referred to the target's race and object and were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 6

Multilevel Linear Regression Predicting Stimulus Anger Ratings from Stimulus Threat Ratings, Target Race, and Target Object in Study 1

	Fixed Effects					
	b	*β*	*SE*	*t*	*df*	*p*
Intercept	41.31	.010	1.65	25.08	25.10	<.001
Threat	.29	.294	.01	125.04	19.43	<.001
Object	-.68	-.026	.43	38.89	-1.58	.123
Race	-.86	-.032	1.48	39.07	-.58	.563
Threat*Race	.03	.028	.01	55.31	2.65	.011

		Random Effects			
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		341.198	3.795	89.91	<.001
Participant	Intercept	120.575	12.502	9.65	<.001
	Covariance (Intercept/Race)	-.571	2.749	-.21	.836
	Race	7.473	1.191	6.27	<.001
	Covariance (Intercept/Threat)	-.184	.142	-1.30	.195
	Covariance (Race/Threat)	.007	.039	.18	.857
	Threat	.016	.003	5.46	<.001
Target	Intercept	85.241	19.7798	4.31	<.001
	Covariance (Intercept/Object)	-1.049	4.253	-.25	.805
	Object	6.718	1.710	3.93	<.001
	Covariance (Intercept/Threat)	.168	.102	1.65	.100
	Covariance (Race/Threat)	-.064	.031	-2.07	.039
	Threat	.003	.001	2.99	.003

Note. N_{sample} = 228. $N_{stimuli}$ = 79. $N_{observations}$ = 17,530. Race and Object were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun). Threat ratings were mean-centered.

Table 7

Descriptive Statistics from Shooter Task for Studies 2 and 3.

	Proportion Errors		Mean Reaction Time	
	Study 2	Study 3	Study 2	Study 3
Overall	.29 (.45)	.26 (.44)	516.80 (71.51)	512.50 (70.16)
Race				
Black	.28 (.45)	.27 (.44)	516.73 (71.85)	511.80 (69.88)
White	.30 (.46)	.26 (.44)	516.87 (71.16)	513.19 (70.42)
Object				
Gun	.26 (.44)	.25 (.43)	502.65 (68.89)	495.44 (67.56)
Non-gun	.32 (.47)	.28 (.45)	534.73 (70.76)	532.96 (67.70)
Race by Object				
Unarmed Black	.32 (.47)	.30 (.46)	533.58 (72.53)	532.79 (67.89)
Unarmed White	.31 (.46)	.26 (.44)	535.90 (68.87)	533.12 (67.53)
Armed Black	.24 (.43)	.24 (.43)	503.56 (68.49)	495.26 (66.93)
Armed White	.28 (.45)	.26 (.44)	501.69 (69.29)	495.63 (68.20)

Note. Standard deviations are presented in parentheses. Reaction time data represent data from correct trials only.

Table 8

Multilevel Logistic Regression Predicting Odds of Error in Study 2, with Perceived Threat Treated as a Property of the Participant

	Fixed Effects				
	b	e^b	*F*	*df*	*p*
Intercept	-1.005	.366			
Threat Bias	.002	1.002	.337	211	.562
Race	-.044	.957	.480	78	.491
Object	-.131	.878	3.784	96	.055
Threat Sum	.002	1.002	2.036	211	.155
Threat Sum x Race	-.000	1.000	.875	35,386	.350
Threat Sum x Object	-.001	.999	3.039	215	.083
Race x Threat Bias	.002	1.002	4.020	35,386	.045
Race x Object	-.055	.947	.744	78	.391
Object x Threat Bias	-.005	.995	7.280	215	.008
Race x Object x Threat Bias	.000	1.000	.126	35,386	.722
Threat Sum x Race x Object	-.000	1.000	1.007	35,386	.316

Random Effects					
Subject	Parameter	Variance	SE	Wald *z*	*p*
Participant	Intercept	.234	.026	8.921	<.001
	Object	.100	.013	7.674	<.001
	Covariance	-.047	.014	-3.442	.001
Stimulus	Intercept	.318	.053	5.993	<.001

Note. N_{sample} = 221. $N_{stimuli}$ = 79. $N_{observations}$ = 35,398. $\chi^2(11) = 146.70$, $p < .001$. Estimates labeled "-.000" were between -.001 and 0, and estimates labeled ".000" were between 0 and .001. Threat Bias: Participant's mean White threat rating subtracted from participant's mean Black threat rating. Threat Sum: Sum of a participant's mean Black threat rating and mean White threat rating. Threat Bias and Threat Sum were mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 9

Multilevel Linear Regression Model Predicting Reaction Time in Study 2, with Perceived Threat Treated as a Property of the Participant

Fixed Effects							
						95% CI for *β*	
	b	*SE*	*β*	*t*	*p*	LB	UB
Intercept	522.06	2.94	522.06	173.83	<.001	516.26	527.86
Threat Bias	-.06	.15	-.70	213.22	.708	-4.40	2.99
Race	-.94	2.33	-.94	79.00	.689	-5.58	3.70
Object	-15.87	2.41	-15.87	90.14	<.001	-20.67	-11.08
Threat Sum	-.08	.06	-2.51	215.05	.176	-6.15	1.14
Race x Threat Sum	-.01	.01	-.26	247.81	.571	-1.17	.64
Object x Threat Sum	.03	.02	.91	206.77	.237	-.61	2.43
Race x Threat Bias	.03	.04	.36	204.64	.430	-.53	1.24
Race x Object	1.07	2.33	1.07	78.28	.646	-3.56	5.70
Object x Threat Bias	-.19	.06	-2.35	197.47	.003	-3.87	-.82
Race x Object x TB	-.03	.03	-.36	20740.82	.393	-1.18	.46
Race x Object x TS	.03	.01	.96	20776.30	.026	.12	1.80

Random Effects					
Subject	Parameter	Variance	*SE*	Wald *z*	*p*
	Residual	3764.54	37.16	101.319	<.001
Stimulus	Intercept	427.10	70.86	6.027	<.001
Participant	Intercept	709.50	72.56	9.778	<.001
	Object	88.25	12.78	6.907	<.001
	Cov. (Int/Obj)	-53.06	22.06	-2.406	.016
	Race	5.68	4.06	1.399	.162
	Cov. (Int/Race)	39.46	13.43	2.939	.003
	Cov. (Obj/Race)	1.45	5.29	.274	.784

Note. N_{sample} = 221. $N_{stimuli}$ = 79. $N_{observations}$ = 21,251. $\chi^2(11) = 27.91$, $p = .003$. Only correct responses were included. Threat Bias (TB): Participant's mean White threat rating subtracted from participant's mean Black threat rating. Threat Sum (TS): Sum of a participant's mean Black threat rating and mean White threat rating. Threat Bias and Threat Sum were mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 10

Multilevel Logistic Regression Predicting Odds of Error in Study 2, with Perceived Threat Treated as a Property of the Stimulus

	Fixed Effects				
	b	e^b	*F*	*df*	*p*
Intercept	-1.009	.365	2.164	75	.047
Race	-.058	.944	.870	73	.354
Object	-.141	.869	4.611	91	.034
Threat	.001	1.001	.022	73	.881
Race x Object	-.062	.940	1.005	73	.319
Object x Threat	.023	1.023	5.942	73	.017
Race x Threat	.015	1.015	2.394	73	.126
Threat x Race x Object	-.002	.998	.043	73	.836

Random Effects

Subject	Parameter	Variance	SE	Wald *z*	*p*
Participant	Intercept	.239	.027	8.958	<.001
	Object	.104	.013	7.777	<.001
	Covariance	-.049	.014	-3.558	<.001
Stimulus	Intercept	.295	.051	5.778	<.001

Note. N_{sample} = 221. $N_{stimuli}$ = 79. $N_{observations}$ = 35,398. $\chi^2(7) = 1831.83$, $p < .001$. Threat: Threat ratings for a given stimulus, summed across participants. Threat was mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1= Gun, -1 = Non-gun).

Table 11

Multilevel Linear Regression Model Predicting Reaction Time in Study 2, with Perceived Threat Treated as a Property of the Stimulus

Fixed Effects							
						95% CI for β	
	b	*SE*	β	*t*	*p*	LB	UB
Intercept	522.11	2.95	522.11	177.14	<.001	516.29	527.93
Race	-1.15	2.34	-1.15	-.49	.626	-5.81	3.52
Object	-15.87	2.43	-15.87	-6.53	<.001	-20.70	-11.04
Threat	.03	.36	.21	.09	.930	-4.58	5.01
Race x Object	1.19	2.34	1.19	.51	.612	-3.47	5.85
Object x Threat	.77	.36	5.14	2.14	.036	.35	9.94
Race x Threat	.35	.36	2.31	.96	.341	-2.49	7.10
Threat x Race x Object	.07	.36	.47	.20	.845	-4.32	5.27

Random Effects					
Subject	Parameter	Variance	*SE*	Wald *z*	*p*
	Residual	3770.70	37.44	100.703	<.001
Stimulus	Intercept	424.09	72.73	5.831	<.001
Participant	Intercept	708.28	72.19	9.812	<.001
	Object	93.76	13.25	7.077	<.001
	Cov. (Int/Obj)	-54.98	22.43	-2.452	.014
	Race	4.18	3.97	3.973	.293
	Cov. (Int/Race)	38.05	13.25	2.873	.004
	Cov. (Obj/Race)	.58	5.32	.110	.912

Note. N_{sample} = 221. $N_{stimuli}$ = 79. $N_{observations}$ = 21,251. $\chi^2(7) = 51.33$, $p < .001$. Only correct responses were included. Threat: Threat ratings for a given stimulus, summed across participants. Threat was mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1= Gun, -1 = Non-gun).

Table 12

Results from a Drift Diffusion Analysis Modeling Error and Reaction Time Data from Study 2

Variable	Modal Estimate	95% HDI LB	95% HDI UB
Intercepts *(Parameter Units)*			
Alpha	1.067	1.048	1.088
Beta	.512	.503	.522
Tau	.949	.947	.950
Delta	.142	.084	.196
Delta *(Transformed Units)*	.036	.021	.049
Slopes for Delta *(Transformed Units)*			
Threat Bias	.020	.005	.034
Threat Sum	.010	-.004	.024
Race	.011	.005	.016
Object	.237	.232	.243
Race x Object	.011	.005	.017
Race x Threat Bias	-.0003	-.006	.005
Race x Threat Sum	.003	-.002	.008
Object x Threat Bias	-.003	-.009	.002
Object x Threat Sum	-.006	-.011	<.001
Race x Object x Threat Bias	-.005	-.011	<.001
Race x Object x Threat Sum	.001	-.004	.007

Note. $N_{observations}$ = 35,398. Threat Bias: Difference between a participant's mean threat ratings for Black targets and for White targets. Threat Sum: Sum of these two means. Both Threat Bias and Threat Sum were scaled and mean-centered to have a mean of zero and a standard deviation equal to 1. Race and Object were effects-coded (1 = Black, -1 = White; 1= Gun, -1 = Non-gun).

Values labeled "Parameter Units" can be interpreted as values of the relevant parameter. Values labeled "Transformed Units" were used to calculate delta values via a transformation involving the distribution function *D*() for a normal distribution with mean 0 and standard deviation 1. The purpose of this transformation was to constrain parameter estimates for delta between -5 and 10 (a generous range of possible values), thus preventing the model from producing impossible values for delta which could lead to estimation problems. Approximate values of delta for a given combination of variable values can therefore be calculated by inserting the values into the expression "-5 + 10**D*(formula)," where the formula uses the "transformed units" delta intercept and slopes listed in the present table.

E.g., the expected value of delta for armed Blacks when Threat Bias and Threat Sum are both one standard deviation above their means (i.e., when all variables = 1) would be approximately:
-5 + 10*D(.036 + .020 + .010 + .011 + .237 + .011 - .0003 + .003 - .003 - .006 -.005 + .001)
= 1.235

Table 13

Multilevel Linear Regression Model Predicting Differences in Participants' Threat Ratings between the Armed and Unarmed Pictures of Each Target in Study 3

	Fixed Effects			
	F	*df 1*	*df 2*	*p*
Intercept	2.736	1	452.193	.099
Target ID	31.954	19	9065.573	<.001
Sum of Threat Ratings	.147	1	1550.854	.702

Random Effects

Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		758.33	11.26	67.334	<.001
Participant	Intercept	20.01	3.63	5.512	<.001

Note. N_{sample} = 233. N_{target} = 20. $N_{observations}$ = 9320.

Table 14

Multilevel Linear Regression Model Predicting Differences in Participants' Threat Ratings between the Two Pictures of Each Target/Object Combination in Study 3

	Fixed Effects			
	F	*df 1*	*df 2*	*p*
Intercept	1499.290	1	231.610	<.001
Target ID	6.195	19	9079.145	<.001
Object	15.382	1	9066.185	<.001
Sum of Threat Ratings	14.485	1	6708.260	<.001

Random Effects

Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		200.84	2.98	67.327	<.001
Participant	Intercept	28.87	3.15	9.164	<.001

Note. N_{sample} = 233. N_{pairs} = 40. $N_{observations}$ = 9320. Object was effects-coded (1 = Gun, -1 = Non-gun).

Table 15

Multilevel Linear Regression Predicting Stimulus-Level Threat Rating in Study 3

	Fixed Effects				
	b	*SE*	*t*	*df*	*p*
Intercept	<.001	1.31	<.001	230.905	1.000
Race	-1.367	1.01	-1.349	109.087	.180
Object	.193	.92	.209	77.000	.835

Random Effects					
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		282.72	2.97	95.121	<.001
Participant	Intercept	202.21	19.10	10.585	<.001
	Covariance	10.45	6.23	1.677	.094
	Race	39.75	4.02	9.891	<.001
Stimulus	Intercept	67.19	11.02	6.095	<.001

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 18,640. Race refers to race of the target. Threat was mean-centered. Race and Object were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 16

Multilevel Linear Regression Model Predicting Stimulus Threat Ratings in Study 3

	Fixed Effects				
	b	*SE*	*t*	*df*	*p*
Intercept	42.84	1.53	28.036	90.845	<.001
Object	.19	.51	.376	38.000	.709
Race	-1.37	1.28	-1.068	47.225	.291
Object*Race	.21	.51	.411	38.000	.683

Random Effects					
Subject	Parameter	Variance	SE	Wald *z*	*p*
Residual		282.72	2.97	95.121	<.001
Participant	Intercept	202.21	19.10	10.585	<.001
	Covariance	10.45	6.23	1.677	.094
	Race	39.75	4.02	9.891	<.001
Target	Intercept	58.10	13.47	4.314	<.001
	Covariance	6.32	4.17	1.516	.129
	Object	9.94	2.42	4.108	<.001

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 18,640. Race and Object referred to the target's race and object and were effects-coded (1 = Black, -1 = White; 1 = Gun, -1 = Non-gun).

Table 17

Multilevel Logistic Regression Predicting Odds of Error in Study 3, with Perceived Threat Treated as a Property of the Participant

	Fixed Effects				
	b	e^b	*F*	*df*	*p*
Intercept	-1.125	.324	.821	133	.000
Threat Bias	-.001	.999	.043	221	.836
Race	.036	1.037	.447	76	.506
Object	-.083	.920	2.122	92	.149
Threat Sum	.000	1.000	.002	278	.967
Threat Sum x Race	-.000	1.000	.723	37,395	.395
Threat Sum x Object	-.001	.999	1.401	252	.238
Race x Threat Bias	-.001	.999	.289	74	.593
Race x Object	.080	.923	2.130	76	.149
Object x Threat Bias	.001	1.001	.209	174	.648
Race x Object x Threat Bias	-.000	1.000	.100	74	.753
Threat Sum x Race x Object	-.001	.999	1.565	37,395	.211

		Random Effects			
Subject	Parameter	Variance	SE	Wald *z*	*p*
Participant	Intercept	.258	.028	9.235	<.001
	Object	-.023	.012	-1.923	.055
	Covariance	.070	.010	7.086	<.001
Stimulus	Intercept	.226	.039	5.833	<.001
	Threat Bias	-.000	.001	-.386	.700
	Covariance	.000	.000	.775	.439

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 37,407. $\chi^2(11) = 156.50$, $p < .001$. Estimates labeled "-.000" were between -.001 and 0, and estimates labeled ".000" were between 0 and .001. Threat Bias: Participant's mean White threat rating subtracted from participant's mean Black threat rating. Threat Sum: Sum of a participant's mean Black threat rating and mean White threat rating.

Table 18

Multilevel Linear Regression Model Predicting Reaction Time in Study 3, with Perceived Threat Treated as a Property of the Participant

Fixed Effects							
						95% CI for β	
	b	*SE*	β	*t*	*p*	LB	UB
Intercept	516.202	2.72	516.202	190.047	<.001	510.840	521.564
Threat Bias	-.074	.13	-.973	-.570	.569	-4.334	2.388
Race	-.009	2.15	-.009	-.004	.997	-4.292	4.275
Object	-17.665	2.21	-17.665	-7.981	<.001	-22.066	-13.264
Threat Sum	-.061	.05	-1.729	-1.159	.247	-4.661	1.203
Race x Threat Sum	-.019	.01	-.537	-1.339	.181	-1.322	.249
Object x Threat Sum	-.027	.02	-.776	-1.197	.232	-2.053	.501
Race x Threat Bias	.013	.03	.166	.382	.704	-.701	1.033
Race x Object	-.968	2.15	-.968	-.450	.654	-5.251	3.316
Object x Threat Bias	-.014	.05	-.178	-.261	.794	-1.527	1.170
Race x Object x TB	-.033	.03	-.435	-.999	.321	-1.303	.432
Race x Object x TS	-.001	.01	-.028	-.069	.945	-.813	.758

Random Effects					
Subject	Parameter	Variance	*SE*	Wald *z*	*p*
	Residual	3646.28	33.97	107.32	<.001
Participant	Intercept	633.93	63.41	10.00	<.001
	Object	61.67	9.45	6.52	<.001
	Covariance	-2.34	17.59	-.13	.894
Stimulus	Intercept	357.34	60.25	5.93	<.001
	Threat Bias	.01	.01	.55	.582
	Covariance	.46	.66	.69	.489

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 23,668. $\chi^2(11) = 25.84$, *p* = .007. Only correct responses were included. Threat Bias (TB): Participant's mean White threat rating subtracted from participant's mean Black threat rating. Threat Sum (TS): Sum of a participant's mean Black threat rating and mean White threat rating.

Table 19

Multilevel Logistic Regression Predicting Odds of Error in Study 3, with Perceived Threat Treated as a Property of the Stimulus

	Fixed Effects				
	b	e^b	*F*	*df*	*p*
Intercept	-1.132	.322			<.001
Race	.024	1.025	.192	72	.662
Object	-.061	.941	1.119	86	.293
Threat	-.006	.994	.827	72	.366
Race x Object	-.073	.930	1.737	72	.192
Object x Threat	.003	1.003	.228	72	.634
Race x Threat	-.002	.998	.112	72	.739
Threat x Race x Object	.015	1.015	4.320	72	.041

Random Effects

Subject	Parameter	Variance	SE	Wald *z*	*p*
Participant	Intercept	.255	.028	9.261	<.001
	Object	.069	.010	7.103	.052
	Covariance	-.023	.012	-1.953	-.046
Stimulus	Intercept	.223	.039	5.672	<.001

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 37,407. $\chi^2(7) = 81.63$, $p < .001$. Threat: Threat ratings for a given stimulus, summed across participants.

Table 20

Multilevel Linear Regression Model Predicting Reaction Time in Study 3, with Perceived Threat Treated as a Property of the Stimulus

Fixed Effects							
						95% CI for β	
	b	*SE*	β	*t*	*p*	LB	UB
Intercept	516.46	2.75	516.46	188.13	<.001	511.03	521.88
Race	-.24	2.19	-.24	-.11	.913	-4.61	4.13
Object	-17.19	2.25	-17.19	-7.63	<.001	-21.68	-12.71
Threat	-.08	.28	-.69	-.30	.768	-5.32	3.95
Race x Object	-.70	2.19	-.70	-.32	.749	-5.08	3.67
Object x Threat	.23	.28	1.86	.80	.425	-2.76	6.49
Race x Threat	.27	.28	2.18	.94	.350	-2.44	6.81
Threat x Race x Object	.38	.28	3.10	1.34	.186	-1.53	7.72

Random Effects					
Subject	Parameter	Variance	*SE*	Wald *z*	*p*
	Residual	3643.83	34.08	106.93	<.001
Stimulus	Intercept	358.31	62.06	5.77	<.001
Participant	Intercept	627.78	62.53	10.04	<.001
	Object	60.75	9.29	6.54	<.001
	Cov. (Int/Obj)	-2.91	17.34	-.17	.867
	Threat	.06	.56	1.05	.296
	Cov. (Int/Threat)	.89	1.36	.66	.511
	Cov. (Obj/Threat)	.06	.51	.12	.905

Note. N_{sample} = 233. $N_{stimuli}$ = 80. $N_{observations}$ = 23,668. $\chi^2(7) = 59.82$, $p < .001$. Only correct responses were included. Threat: Threat ratings for a given stimulus, summed across participants.

Table 21

Results from a Drift Diffusion Analysis Modeling Error and Reaction Time Data from Study 3

Variable	Modal Estimate	95% HDI LB	95% HDI UB
Intercepts *(Parameter Units)*			
Alpha	1.072	1.054	1.092
Beta	.518	.508	.527
Tau	.948	.947	.950
Delta	.090	.030	.145
Delta *(Transformed Units)*	.023	.008	.036
Slopes for Delta *(Transformed Units)*			
Threat Bias	-.009	-.022	.007
Threat Sum	.008	-.006	.022
Race	.017	.011	.022
Object	.272	.266	.278
Race x Object	-.004	-.009	.002
Race x Threat Bias	.002	-.003	.007
Race x Threat Sum	.004	-.002	.009
Object x Threat Bias	.007	.001	.012
Object x Threat Sum	.004	-.001	.010
Race x Object x Threat Bias	.002	-.004	.006
Race x Object x Threat Sum	.002	-.004	.007

Note. $N_{observations}$ = 37,407. Threat Bias: Difference between a participant's mean threat rating for Black targets and their mean threat rating for White targets. Threat Sum: Sum of these two mean ratings. Threat Bias and Threat Sum were scaled and mean-centered so as to have a mean of zero and a standard deviation equal to 1. Race and Object were effects-coded (1 = Black, -1 = White; 1= Gun, -1 = Non-gun).

Values labeled "Parameter Units" can be interpreted as values of the relevant parameter. Values labeled "Transformed Units" were used to calculate delta values via a transformation involving the distribution function *D*() for a normal distribution with mean 0 and standard deviation 1. The purpose of this transformation was to constrain parameter estimates for delta between -5 and 10 (a generous range of possible values), thus preventing the model from producing impossible values for delta which could lead to estimation problems. Approximate values of delta for a given combination of variable values can therefore be calculated by inserting the values into the expression "-5 + 10**D*(formula)," where the formula uses the "transformed units" delta intercept and slopes listed in the present table.

E.g., the expected value of delta for armed Blacks when Threat Bias and Threat Sum are both one standard deviation above their means (i.e., when all variables = 1) would be approximately:
$-5 + 10*D(.023 - .009 + .008 + .017 + .272 - .004 + .002 + .004 + .007 + .004 + .002 + .002)$
$= 1.285$

APPENDIX B: FIGURES

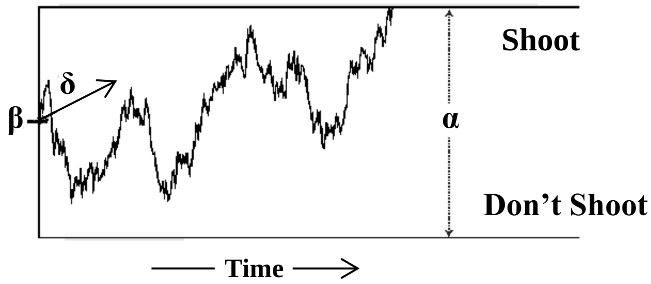

Figure 1. Evidence accumulation in the shooting decision process according to the Drift Diffusion Model.

Figure 2. Example targets for the First-Person Shooting Task.

Stimuli in the task will not be labeled "Experimental Stimulus"; this is included here to comply with IRB stipulations about non-experimental uses of these individuals' pictures.

APPENDIX C: POWER SIMULATIONS

Table C1

Equation for White Threat Ratings

Coefficient	Term	Explanation
0.034	(intercept)	Intercept from pilot study
0.700	Black threat ratings	Slope from pilot study
0	Version of survey	No hypothesized effect
0.015	Black threat ratings X version	Version of survey causes a difference in standardized slopes of 0.03
N~(0, 0.76)	Error	Residual standard error from pilot was 0.7576

Note. Equation predicting White threat ratings used in power simulations for Study 1.

Table C2

Equation for Reaction Times for Participant-Level Threat

Coefficient	Term	Explanation
6.22	Intercept	Intercept from Study 1 of my master's thesis
-0.0237	Object	Slope from Study 1 of master's thesis
0.001	Race	Slope found in both master's thesis and Correll et al. (2002)
0	Threat Bias (difference B-W)	No hypothesized effect
-0.001	Threat Sum (B+W)	Hypothesized small effect: greater perceived threat may be associated with faster responding.
-0.015	Object X Race	race:object interaction term from Correll et al. (2002), where this term was significant
-0.001	Threat Bias X Race	Hypothesized small effect: may respond slightly faster to Blacks if perceive Blacks as more threatening
0	Threat Bias X Object	No hypothesized effect
0	Threat Sum X Race	No hypothesized effect
-0.001	Threat Sum X Object	Hypothesized small effect: may respond faster to guns if greater overall threat perception
-0.008	Object X Threat Bias X Race	Hypothesized that 1 standard deviation increase in threat bias leads to increase in shooter bias; magnitude equivalent to the standard error of the Correll et al. (2002) race:object interaction term
0	Object X Threat Sum X Race	No hypothesized effect
N~(0, .099)	Participant-level random intercepts	0.099 = Standard deviation of random intercepts for participants in first master's study (when using target-level slopes for object)

Table C2 (cont'd)

N~(0, .016)	Target-level random intercepts	0.016 = Standard deviation of random intercepts for targets in first master's study (when using target-level slopes for object)
N~(0, .02)	Target-level random slopes for threat bias	0.02 = Standard deviation of random object slopes for targets in first master's study (when using target-level slopes for object)
N~(0, .02)	Target-level random slopes for threat bias	0.02 = Standard deviation of random object slopes for targets in first master's study (when using target-level slopes for object)
N~(0, .23)	Error	0.23 = Standard deviation of residual random effects from first master's study.

Note. Equation generating log-transformed reaction times used in power simulations for Study 2 analysis treating threat bias as a participant-level variable.

Table C3

Equation for Reaction Times for Stimulus-Level Threat

Coefficient	Term	Explanation
6.220	Intercept	Intercept from Study 1 of my master's thesis
-0.0237	Object	Slope from Study 1 of master's thesis
0.001	Race	Slope found in both master's thesis and Correll et al. (2002)
-0.001	Target threat rating	Hypothesized small effect: targets perceived as more threatening may receive faster responses
-0.015	Object X Race	race:object interaction term from Correll et al. (2002), where this term was significant
-0.001	Target threat rating X Race	Hypothesized small effect: Black people who are perceived as more threatening may receive faster responses
-0.015	Object X Target threat rating	Shoot responses to targets perceived as threatening may be faster than shoot responses to targets perceived as non-threatening. (This value copies the object:race coefficient. All variables are standardized to have mean = 1 and *SD* = 1)
-0.008	Object X Target threat rating X Race	Shooter bias may be more relevant to targets perceived as more threatening. Coefficient here equivalent to the standard error of the Correll et al., (2002) race:object interaction term
N~(0, .1)	Participant-level random intercepts	0.1 = Standard deviation of random intercepts for participants in first master's study (when using participant-level slopes for object rather than target-level slopes)

Table C3 (cont'd)

N~(0, .036)	Participant-level random slopes for target threat ratings	0.036 = Standard deviation of random object slopes for participants in first master's study (when using participant-level slopes for object rather than target-level slopes)
N~(0, .017)	Target-level random intercepts	0.017 = Standard deviation of random intercepts for targets in first master's study (when using participant-level slopes for object rather than target-level slopes)
N~(0, .23)	Error	0.23 = Standard deviation of residual random effects from first master's study.

Note. Equation generating reaction times used in power simulations for Study 2 analysis treating threat as a target-level variable.

R script for simulations.

```
#To maximize readability of this script, it is recommended that the reader
view it in RStudio. However, it is included in this manuscript in order to
provide a comprehensive justification of the proposed research.

library(truncnorm)
library(car)
library(lme4)
library(paramtest)
library(pwr)
library(ggplot2)
library(knitr)
library(lavaan)
library(dplyr)
```

Study 1.

```
#Test of Whether Survey Version Interacts with  Black Threat Ratings as
Predictor of White Threat Ratings####

sim_mod_s1 <- function(simNum, N){ #function to simulate data

  #randomly generate independent variables

  pnum = 1:N #participant indicator variable

  #participants' mean threat ratings
  #for Black targets:
  bthreat = truncnorm::rtruncnorm(N, mean = 43.35, a = 0, b = 100, sd =
13.87) #values from pilot

  #version of the survey which participants completed
  version <- rep(c(-1,1), each = (N/2))

  #initiate dataset
  fakedata <- as.data.frame(cbind(pnum, bthreat, version))

  #scale Black threat ratings to have mean = 0 and SD = 1
  fakedata$bthreat <- scale(fakedata$bthreat)

  #create DV: mean White target ratings
  fakedata$wthreat <- NA
  fakedata$wthreat <- 0.034 + #intercept from pilot study
    0.70*bthreat +  #main effect of Black threat ratings similar to pilot
effect
    0*version + #no hypothesized main effect of version
    .015*bthreat*version + #version of survey (condition) causes a difference
in standardized slopes of 0.03
    rnorm(N, mean = 0, sd = 0.76) #residual standard error from pilot was
0.7576.

  #scale White threat ratings to have mean = 0 and SD = 1
  fakedata$wthreat <- scale(fakedata$wthreat)
```

```
  returnP <- tryCatch({
    mod1 <- lm(wthreat ~ bthreat*version, data = fakedata)
    out <- summary(mod1)
    p <- out$coefficients[4,4]
    sig <- (p<.05)
    return(c(p,sig)) #returns p-value and True/False indicator of whether it
was significant
  },
  error = function(e){ #if model throws an error, we'll get "NA" for that run
instead of a p-value
    return(c(p = NA, sig = NA))
  })

  return(returnP)

}

power_sim_mod_s1 <- grid_search(sim_mod_s1,
                              params=list(N=c(200)), #list of sample sizes of
interest
                              n.iter=500, #run the simulation n.iter times
per sample size listed above
                              output='data.frame', parallel='snow',ncpus=4)

#output is a chart indicating proportion of significant (non-NA) results for
each sample size (=power), and number of NA results
results(power_sim_mod_s1) %>%
  group_by(N.test) %>%
  summarize(
    power=mean(X2, na.rm=T),
    na=sum(is.na(X2))
  )
```

Study 2.

```
#Explanation of references to studies:
  #"masters s1" is shorthand for the first study in my master's thesis. I
sometimes draw values from this study's output to inform my simulation.
  #"Correll 2002" refers to a study by Correll et al. (2002) which had a
significant race:object effect for reaction time data when random effects
were controlled for. I sometimes draw values from this study as well.
```

Treating threat bias as a property of the participant.

```
#Reaction Time: Threat bias as a participant variable####
shooter_mod_rt_P <- function(simNum, N){  #function to simulate data

  #randomly generate independent variables and grouping variables
  pnum = rep(1:N, each = 160) #participant indicator variable

  #participants' mean threat ratings
    #for Black targets:
```

```
  pbthreat = truncnorm::rtruncnorm(N, mean = 43.35, a = 0, b = 100, sd = 
13.87) #values from pilot
  pbthreatlong = rep(pbthreat, each = 160)
    #for White targets:
  pwthreat = truncnorm::rtruncnorm(N, mean = 41.32, a = 0, b = 100, sd = 
11.15 ) #values from pilot
  pwthreatlong = rep(pwthreat, each = 160)

  #target indicator variable
  targetgp = rep(1:40, each = 4)
  target = sample(targetgp, size = 160, replace = F)

  #create data set called fakedata:
  fakedata <- as.data.frame(cbind(pnum,pwthreatlong, pbthreatlong))
  fakedata$target <- NA
  fakedata$race <- NA
  fakedata$obj <- NA
  for(i in(unique(pnum))){
    fakedata$target[which(fakedata$pnum==i)] <- sample(targetgp, size = 160, 
replace = F)
  }
  #generating some race and object values for each trial:
  for(i in(1:nrow(fakedata))){
    fakedata$race[i] <- ifelse(fakedata$target[i]<11, -1, 1) #1 is Black; -1 
is White. Targets numbered 1-10 are White; targets numbered 11-20 are Black
    #80 trials with each object for each participant:
    objsofar <- fakedata$obj[which(fakedata$target==fakedata$target[i] & 
fakedata$pnum == fakedata$pnum[i])]
    if(length(which(!(is.na(objsofar))))<2) {
      fakedata$obj[i] <- -1 #-1 is nongun
    }
    else {fakedata$obj[i] <- 1} #1 is gun
  }

  #randomly designating some trials as Errors so we can exclude those as we 
would in a real reaction time analysis, limiting the number of included 
trials
  gunvec <- c(1,1,1,1,1,1,1,-1,-1,-1) #in masters s1, participants shot on 
about 70% of gun trials
  nongunvec<- c(1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1) #in 
masters s1, participants shot on about 35% of nongun trials
  fakedata$shoot <- NA
  fakedata$shoot[which(fakedata$obj==1)] <- sample(gunvec) #picks ~70% of gun 
trials where the participant chose "shoot"
  fakedata$shoot[which(fakedata$obj==-1)] <- sample(nongunvec) #picks ~35% of 
nongun trials where the participant chose "shoot"
  fakedata$Error <- ifelse(fakedata$shoot==1,
                          ifelse(fakedata$obj==1,0,1), #if they chose 
"shoot," correct (0) for guns and incorrect (1) for nonguns
                          ifelse(fakedata$obj==-1,0,1) ) #if they chose 
"don't shoot," correct (0) for nonguns and incorrect (1) for guns

  #participant-level threat variables
  fakedata$pbias <- fakedata$pbthreatlong-fakedata$pwthreatlong #difference 
score for participant-level mean threat ratings: Black minus White
  fakedata$psum <- fakedata$pbthreatlong+fakedata$pwthreatlong #sum of 
participant-level mean threat ratings: Black plus White
```

```
  #these variables need to be factors
  fakedata$target <- as.factor(fakedata$target)
  fakedata$pnum <- as.factor(fakedata$pnum)

  #generating random values for the random slopes and intercepts
    #participant intercept:
  pnum_int <- rep(rnorm(N, 0, 0.099), each = 160) #standard deviation of
random intercepts for participants in masters s1
    #target intercept:
  t_int <- rep(rnorm(N, 0, .016), each = 160) #standard deviation of  random
intercepts for targets in masters s1
    #target slopes for threat bias (difference score):
  t_slope_diff <- rep(rnorm(N, 0, .02), each = 160) # standard deviation of
random object slopes for targets in masters s1
    #target slopes for sum of participants' mean Black and White threat
ratings:
  t_slope_sum <- rep(rnorm(N, 0, .02), each = 160) # standard deviation of
random object slopes for targets in masters s1

  #scale the participant-level threat variables (mean-center and set SD to 1)
  fakedata$pbias <- scale(fakedata$pbias)
  fakedata$psum <- scale(fakedata$psum)

  #generate DV from the data created above, plus some randomness
  fakedata$logRT <- (6.22 + pnum_int + t_int) + #intercept from master s1
    (-0.0237)*fakedata$obj + #object slope from master s1
    0.001*fakedata$race + #race slope from master s1 AND correll 2002
    (t_slope_diff)*fakedata$pbias + #no hypothesized effect
    (-0.001 + t_slope_sum)*fakedata$psum + #hypothesized small effect:
greater perceived threat may be assocd with faster responding.
    (-0.015)*fakedata$obj*fakedata$race + #race:object interaction term from
Correll et al., 2002, where this term was significant. Negative coefficient
means bias toward shooting armed people faster if Black, and don't-shooting
unarmed people faster if White
    -0.001*fakedata$pbias*fakedata$race + #hypothesized small effect: may
respond slightly faster to Blacks if perceive Blacks as more threatening
    0*fakedata$obj*fakedata$pbias + #no hypothesized effect
    0*fakedata$psum*fakedata$race + #no hypothesized effect
    -0.001*fakedata$obj*fakedata$psum + #hypothesized small effect: may
respond faster to guns if greater overall threat perception
    (-0.008)*fakedata$obj*fakedata$pbias*fakedata$race + #hypothesized: 1
standard dev increase in threat bias leads to increase in shooter bias;
magnitude equivalent to the standard error of the correll 2002 race:obj
interaction term
    0*fakedata$obj*fakedata$psum*fakedata$race + #no hypothesized effect
    rnorm(N*160, 0,0.23) #standard dev of residual random effects from master
s1

  #create models with and without the three-way interaction (threat bias
interacting with shooter bias)
  #then compare to get significance of that interaction
  #(must do model comparison because this function doesn't report p-values)
  returnP <- tryCatch({
    mod1 <- lme4::lmer(logRT ~ race*obj*pbias + race*obj*psum + (1|pnum) +
(1+pbias+psum|target), data = fakedata[which(fakedata$Error==0),])
```

```
    mod2 <- lme4::lmer(logRT ~ race*obj*pbias + race*obj*psum -
race:obj:pbias + (1|pnum) + (1 + pbias + psum|target), data =
fakedata[which(fakedata$Error==0),])
    out <- anova(mod2,mod1)
    p <- out[8][2,1]
    sig <- (p<.05)
    return(c(p,sig)) #returns p-value and True/False indicator of whether it
was significant
  },
  error = function(e){ #if models throw an error (e.g. if they don't
converge), we'll get "NA" for that run instead of a p-value
    return(c(p = NA, sig = NA))
  })

  return(returnP)
}


#Now that the function above has been created, we run the simulation many
times for each sample size of interest
starttime <- proc.time()
power_sim_rt_p <- grid_search(shooter_mod_rt_P,
                              params=list(N=c(200)), #list of sample sizes of
interest
                              n.iter=500, #run the simulation n.iter times
per sample size listed above
                              output='data.frame', parallel='snow',ncpus=4)

#output is a chart indicating proportion of significant (non-NA) results for
each sample size (=power), and number of NA results
results(power_sim_rt_p) %>%
  group_by(N.test) %>%
  summarize(
    power=mean(X2, na.rm=T),
    na=sum(is.na(X2))
  )
endtime<-proc.time()
endtime-starttime #indicates how long the simulations took
```

Treating threat as a property of the target.

```
#Reaction Time:  Threat as a target variable####
shooter_mod <- function(simNum, N){  #function to simulate data

  #randomly generate independent variables and grouping variables
  pnum = rep(1:N, each = 160) #participant indicator variable

  #target indicator variable
  targetgp = rep(1:40, each = 4)
  target = sample(targetgp, size = 160, replace = F)

  #create data set called fakedata:
  fakedata<-as.data.frame(pnum)
  fakedata$target <- NA
```

```
  fakedata$race <- NA
  fakedata$obj <- NA
  for(i in(unique(pnum))){
    fakedata$target[which(fakedata$pnum==i)] <- sample(targetgp, size = 160,
replace = F)
  }
  #generating some race and object values for each trial:
  for(i in(1:nrow(fakedata))){
    fakedata$race[i] <- ifelse(fakedata$target[i]<11, -1, 1) #1 is Black; -1
is White. Targets numbered 1-10 are White; targets numbered 11-20 are Black
    objsofar <- fakedata$obj[which(fakedata$target==fakedata$target[i] &
fakedata$pnum == fakedata$pnum[i])]
    if(length(which(!(is.na(objsofar))))<2) {
      fakedata$obj[i] <- -1 #-1 = nongun
    }
    else {fakedata$obj[i] <- 1} #1 = gun
  }

  #randomly designating some trials as Errors so we can exclude those as we
would in a real reaction time analysis, limiting the number of included
trials
  gunvec <- c(1,1,1,1,1,1,1,-1,-1,-1) #in masters s1, participants shot on
about 70% of gun trials
  nongunvec<- c(1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1) #in
masters s1, participants shot on about 35% of nongun trials
  fakedata$shoot <- NA
  fakedata$shoot[which(fakedata$obj==1)] <- sample(gunvec) #picks ~70% of gun
trials where the participant chose "shoot"
  fakedata$shoot[which(fakedata$obj==-1)] <- sample(nongunvec) #picks ~35% of
nongun trials where the participant chose "shoot"
  fakedata$Error <- ifelse(fakedata$shoot==1,
                          ifelse(fakedata$obj==1,0,1), #if they chose
"shoot," correct (0) for guns and incorrect (1) for nonguns
                          ifelse(fakedata$obj==-1,0,1) ) #if they chose
"don't shoot," correct (0) for nonguns and incorrect (1) for guns

  #target threat values--generate random "mean threat rating" for each target
based on distribution observed in pilot study
  threat <- c(
    truncnorm::rtruncnorm(20, a = 0, b=100, mean = 41.31, sd = 28.12), #white
    truncnorm::rtruncnorm(20, a = 0, b=100, mean=42.59, sd = 25.84) #black
  )
  fakedata$threat <- NA
  for(i in(1:40)){
    fakedata$threat[which(fakedata$target==i)] <- threat[i]
  }

  #these grouping variables need to be factors
  fakedata$target <- as.factor(fakedata$target)
  fakedata$pnum <- as.factor(fakedata$pnum)

  #generating random values for the random slopes and intercepts
    #participant intercepts:
  pnum_int <- rep(rnorm(N, 0, 0.100), each = 160) #standard deviation of
random intercepts for participants in masters s1 (when using participant
slopes rather than object slopes)
    #participant slopes for target threat ratings
```

```
  p_slope <- rep(rnorm(N, 0, .036), each = 160) #standard deviation of random
object slopes for participants in masters s1 (when using participant slopes
rather than object slopes)
    #target intercepts:
  t_int <- rnorm(40, 0, .017)  #standard deviation of  random intercepts for
targets in masters s1 (when using participant slopes rather than object
slopes)
  fakedata$t_int <- NA
  for(i in(1:40)){
    fakedata$t_int[which(fakedata$target==i)] <- t_int[i]
  }

    #scale the target-level threat variable (mean-center and set SD to 1)
  fakedata$threat <- scale(fakedata$threat)

  #generate DV from the data created above, plus some randomness
  fakedata$logRT <- (6.22 + pnum_int + fakedata$t_int) + #intercept from
master s1
    (-0.0237)*fakedata$obj + #object slope from master s1
    0.001*fakedata$race + #race slope from master s1 AND correll 2002
    (-0.001 + p_slope)*fakedata$threat + #hypothesized small effect: targets
perceived as more threatening may receive faster responses
    (-0.015)*fakedata$obj*fakedata$race + #race:object interaction term from
Correll et al., 2002, where this term was significant. Negative coefficient
means bias toward shooting armed people faster if Black, and don't-shooting
unarmed people faster if White
    -0.001*fakedata$threat*fakedata$race + #hypothesized small effect: Black
people who are perceived as more threatening may receive faster responses
    -0.015*fakedata$obj*fakedata$threat + #(this value copies the obj:race
coefficient 2 lines above) Shoot responses to targets perceived as
threatening may be faster than shoot responses to targets perceived as non-
threatening
    -0.008*fakedata$obj*fakedata$threat*fakedata$race + #shooter bias may be
more relevant to targets perceived as more threatening. Magnitude equivalent
to the standard error of the correll 2002 race:obj interaction term
    rnorm(N*160, 0,0.23) #standard dev of residual random effects from master
s1

  #create models with and without the three-way interaction (threat bias
interacting with shooter bias)
  #then compare to get significance of that interaction
  #(must do model comparison because this function doesn't report p-values)
  returnP <- tryCatch({
    mod1 <- lme4::lmer(logRT ~ race*obj*threat + (1+threat|pnum) +
(1|target), data = fakedata[which(fakedata$Error==0),])
    mod2 <- lme4::lmer(logRT ~
race+obj+threat+race:obj+race:threat+obj:threat + (1+threat|pnum) +
(1|target), data = fakedata[which(fakedata$Error==0),])
    out <- anova(mod2,mod1)
    p <- out[8][2,1]
    sig <- (p<.05)
    return(c(p,sig)) #returns p-value and True/False indicator of whether it
was significant
  },
  error = function(e){ #if models throw an error (e.g. if they don't
converge), we'll get "NA" for that run instead of a p-value
    return(c(p = NA, sig = NA))
```

```
  })

  return(returnP)
}

#Now that the function above has been created, we run the simulation many
times for each sample size of interest
starttime <- proc.time()
power_sim <- grid_search(shooter_mod,
                         params=list(N=c(200)), #list of sample sizes of
interest
                         n.iter=500, #run the simulation n.iter times per
sample size listed above
                         output='data.frame', parallel='snow',ncpus=4)

#output is a chart indicating proportion of significant (non-NA) results for
each sample size (=power), and number of NA results
results(power_sim) %>%
  group_by(N.test) %>%
  summarize(
    power=mean(X2, na.rm=T),
    na=sum(is.na(X2))
  )
endtime<-proc.time()
endtime-starttime #indicates how long the simulations took
```

REFERENCES

REFERENCES

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., ... & Schaller, M. (2006). They all look the same to me (unless they're angry) from out-group homogeneity to out-group heterogeneity. *Psychological Science*, *17*(10), 836-840.

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single-and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, *140*(4), 637.

Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review*, *77*(1), 32.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307.

Carré, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*(1651), 2651-2656.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20*(10), 1194-1198.

Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic?. *Perception*, *39*(3), 356-377.

Cooper, R. M., & Langton, S. R. (2006). Attentional bias to angry faces using the dot-probe task? It depends when you look for it. *Behaviour Research and Therapy*, *44*(9), 1321-1329.

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.

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.

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.

Correll, J., Wittenbrink, B., Crawford, M. T., & Sadler, M. S. (2015). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology, 108*(2), 219.

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.

Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, *42*(1), 120-128.

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of clinical and experimental neuropsychology*, *24*(2), 200-213.

Doise, W., Csepeli, G., Dann, H. D., Gouge, C., Larsen, K., & Ostell, A. (1972). An experimental investigation into the formation of intergroup representations. *European Journal of Social Psychology, 2,* 202-204.

Dunham, Y. (2011). An angry = outgroup effect. *Journal of Experimental Social Psychology*, *47*(3), 668-671.

Ember, C. R. (1978). Myths about hunter-gatherers. *Ethnology*, *17*(4), 439-448.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion*, *14*(1), 61-92.

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology*, *44*(1), 164-172.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of Personality and Social Psychology*, *54*(6), 917.

Harder, J. A. (in press). The multiverse of methods: Extending the multiverse analysis to address data collection decisions. *Perspectives on Psychological Science.*

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*, *6*(2), 193.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*(6), 640-643.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, *15*(5), 342-345.

Hunsinger, M. (2011). Threat on the mind: The impact of incidental fear on race bias in rapid decision-making. D*issertation Abstracts International: Section B: The Sciences and Engineering*, *72*(1-B), 594.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*(4), 616.

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.

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.

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.

Kenworthy, J. B., Barden, M. A., Diamond, S., & del Carmen, A. (2011). Ingroup identification as a moderator of racial bias in a shoot–no shoot decision task. *Group Processes & Intergroup Relations*, *14*(3), 311-318.

Kohler, C. G., Turner, T., Stolar, N. M., Bilker, W. B., Brensinger, C. M., Gur, R. E., & Gur, R. C. (2004). Differences in facial expressions of four universal emotions. *Psychiatry Research*, *128*(3), 235-244.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387-15392.

Lundqvist, D., Esteves, F., & Ohman, A. (1999). The face of wrath: Critical features for conveying facial threat. *Cognition & Emotion*, *13*(6), 691-711.

Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to shoot. *Journal of Experimental Social Psychology*, *47*(2), 391-396.

Ma, D. S., Correll, J., Wittenbrink, B., Bar-Anan, Y., Sriram, N., & Nosek, B. A. (2013). When fatigue turns deadly: The association between fatigue and racial bias in the decision to shoot. *Basic and Applied Social Psychology*, *35*(6), 515-524.

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.

Mekawi, Y., Bresin, K., & Hunter, C. D. (2016). White fear, dehumanization, and low empathy: Lethal combinations for shooting biases. *Cultural Diversity and Ethnic Minority Psychology*, *22*(3), 322–332.

Miller, S. L., Zielaskowski, K., & Plant, E. A. (2012). The basis of shooter biases beyond cultural stereotypes. *Personality and Social Psychology Bulletin*, *38*(10), 1358-1366.

Musolino, E. (2012). *The influence of target gender on shooting behaviour: An examination of the role of automatic bias and controlled processes* (Doctoral dissertation, Carleton University, Ottawa, Ontario, Canada). Retrieved from https://curve.carleton.ca/

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*(3), 381.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Park, S. H., & Glaser, J. (2011). Implicit motivation to control prejudice and exposure to counterstereotypic instances reduce spontaneous discriminatory behavior. *Korean Journal of Social and Personality Psychology, 25*(4), 107-20.

Park, S. H., & Kim, H. J. (2015). Assumed race moderates spontaneous racial bias in a computer-based police simulation. *Asian Journal of Social Psychology*, *18*(3), 252-257.

Payne, B. K. (2001). Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*(2), 181.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, *38*(4), 384-396.

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*(1), 195-203.

Perry, R., & Sibley, C. G. (2012). Big-Five personality prospectively predicts social dominance orientation and right-wing authoritarianism. *Personality and Individual Differences*, *52*(1), 3-8.

Pietraszewski, D., & Schwartz, A. (2014). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior, 35*(1), 51-57.

Plant, E. a., Goplen, J., & Kunstman, J. W. (2011). Selective responses to threat: The roles of race and gender in decisions to shoot. *Personality and Social Psychology Bulletin, 37,* 1274-1281.

Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, *16*(3), 180-183.

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology, 41*(2), 141-156.

Pleskac, T. J., Cesario, J., & Johnson, D. J. (2018). How race affects evidence accumulation during the decision to shoot. *Psychonomic Bulletin & Review, 25*(4), 1301-1330.

Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* (Vol. 124, No. 125.10, pp. 1-10).

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873-922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347-356.

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.

Shasteen, J. R., Sasson, N. J., & Pinkham, A. E. (2015). A detection advantage for facial threat in the absence of anger. *Emotion*, *15*(6), 837.

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.

Snowden, A. K. (2017). *Induced Emotions on Shoot Decisions* (Doctoral dissertation, University of Texas-Arlington). Retrieved from https://rc.library.uta.edu/uta-ir/handle/10106/26759

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Stillman, T. F., Maner, J. K., & Baumeister, R. F. (2010). A thin slice of violence: Distinguishing violent from nonviolent sex offenders at a glance. *Evolution and Human Behavior*, *31*(4), 298-303.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149-178.

Taylor, A. C. (2011). *The influence of target race on split-second shooting decisions in simulated scenarios: A Canadian perspective* (Doctoral dissertation, Carleton University, Ottawa, Ontario, Canada). Retrieved from https://curve.carleton.ca/

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*(7), 778.

Todd, A. R., Johnson, D. J., Lassetter, B., Neel, R., Simpson, A. J., & Cesario, J. (2019). *Category salience and racial bias in weapon identification: A diffusion modeling approach.* Manuscript submitted for publication.

Trawalter, S., Todd, A. R., Baird, A. A., & Richeson, J. A. (2008). Attending to threat: Race-based patterns of selective attention. *Journal of Experimental Social Psychology*, *44*(5), 1322-1327.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206-1220.

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.

Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, 116-132.