Check for updates

European Journal of Social Psychology

WILEY
EASP

RESEARCH ARTICLE OPEN ACCESS

The Nature of Racial Bias in the Weapon Identification Task: Discrimination Bias, Response Bias, or Both?

Ruben Laukenmann | Edgar Erdfelder

University of Mannheim, Mannheim, Germany

Correspondence: Ruben Laukenmann (ruben.laukenmann@unisg.ch)

Received: 16 February 2026 | **Revised:** 10 August 2026 | **Accepted:** 17 September 2026

Keywords: discrimination bias | multinomial processing tree modeling | process dissociation procedure | racial profiling | response bias | weapon identification task

ABSTRACT

In the Weapon Identification Task (WIT), participants more often misclassify innocuous objects (e.g., tools) as weapons after seeing a Black compared to a White male face prime. Racially biased object discrimination (discrimination bias) and biased response tendencies (response bias) have been proposed as cognitive mechanisms underlying this weapon identification bias. In three experiments (total $N = 1.015$), we manipulated base-rates of target objects in the WIT and used a generalized process dissociation model to disentangle discrimination bias and response bias. Our results show that response bias is the major source of racial bias in the WIT, while discrimination bias plays a role under specific circumstances only. Specifically, when participants are explicitly instructed to use racial profiling for responding, discrimination bias comes into play in addition to response bias. This suggests that object discrimination is systematically distorted in the WIT whenever racial profiling is used as a response strategy.

1 | Introduction

If an object is presented along with a person, the person's race can influence how well people identify this object (Payne 2001; Payne and Correll 2020; Rivers 2017). One of the most prominent procedures to investigate the influence of race on identifying potentially dangerous objects is the Weapon Identification Task (WIT) (Payne 2001). The WIT is a sequential priming procedure in which a Black or a White male face prime precedes an object. Participants are instructed to quickly identify this object as either a weapon (i.e., a gun) or an innocuous object (i.e., a tool). Results robustly show that innocuous objects are more often misidentified as weapons when preceded by a Black rather than a White face prime, a pattern known as weapon identification bias (Payne 2001; Rivers 2017; Todd et al. 2021).

Understanding the cognitive mechanisms underlying this weapon identification bias has important practical implications. The WIT was originally developed to investigate rapid decisions police officers make when determining whether a suspect is armed or not (Payne 2001). Obviously, errors in such judgements can have severe consequences. Mistakenly perceiving an unarmed suspect as armed may contribute to fatal police encounters (especially for Black males), whereas failing to recognize an armed suspect may place police officers at substantial risk of injury or even death (Edwards et al. 2019; Sierra-Arévalo and Nix 2020; Sierra-Arévalo 2021). Moreover, racially biased policing undermines public trust and citizens' willingness to cooperate with law enforcement (Desmond et al. 2016). Consequently, understanding the cognitive mechanisms contributing to weapon identification bias is important for developing interventions and police training that minimize racial bias while maintaining decision accuracy (Payne et al. 2002; Plant and Peruche 2005; Lai and Lisnek 2023).

Beyond its practical relevance, weapon identification bias raises a fundamental theoretical question regarding the cognitive mechanisms that produce this behavioral pattern. Behavioral

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *European Journal of Social Psychology* published by John Wiley & Sons Ltd.

data alone cannot uncover these mechanisms because identical response patterns may arise from different combinations of latent cognitive components (Gonsalkorale et al. 2009; Jacoby 1991; Klauer et al. 2015; Klein and Todd 2024; Payne 2001; Payne et al. 2002; Todd et al. 2021). Disentangling underlying cognitive components is important not only for theoretical but also for practical reasons, because interventions that successfully reduce one source of racial bias may be ineffective when racial bias originates from another source (Calanchini et al. 2021; Lai et al. 2014; Röhner and Lai 2021). For example, effective interventions will likely differ depending on whether racial bias reflects biased target object perception, stereotype-driven response tendencies, or deliberate profiling strategies (Calanchini et al. 2021; Lai et al. 2014; Paluck et al. 2021; Payne et al. 2002; Plant and Peruche 2005).

Parameterized cognitive models address this ambiguity by decomposing observed responses into measures of latent cognitive components that correspond to theoretically distinct psychological mechanisms (Erdfelder et al. 2009; Klauer and Voss 2008; Klauer et al. 2015; Klein and Todd 2024; Laukenmann et al. 2023; Todd et al. 2021). Rather than inferring mechanisms from overall response patterns, such parameterized cognitive models permit competing mechanistic explanations to be evaluated directly. Consequently, they provide a stronger basis for assessing and explaining how racial bias emerges in rapid decisions and for evaluating experimental manipulations and response strategy training that might modify the underlying cognitive mechanisms (Calanchini et al. 2021; Lai et al. 2014; Röhner and Lai 2021; Todd et al. 2021).

However, interpreting model parameters as measures of cognitive mechanisms presupposes that the respective model provides an adequate representation of the cognitive components underlying response patterns in the WIT (Erdfelder et al. 2009; Klauer et al. 2015). Successfully testing the model's fit to observed data and evaluating whether a cognitive model adequately represents the cognitive mechanisms underlying performance in the WIT is therefore a prerequisite for drawing theoretical conclusions about the origins of weapon identification bias and for evaluating how instructions or training may alter its underlying cognitive mechanisms.

One potential response strategy in the WIT context is racial profiling, that is, the deliberate use of racial stereotypes as a cue for judging whether a person poses a threat. Again, examining effects of racial profiling in the WIT is important for both practical and theoretical reasons. From a practical perspective, racial profiling remains a highly contested policing practice that has received sustained attention in both legal and psychological research because race continues to influence police stop, search, and threat-related decisions (Glaser 2015; Pierson et al. 2020). Hence, determining how deliberately relying on race stereotypes amplifies racial bias is therefore directly relevant for evaluating the consequences of racial profiling as a policing strategy (Payne et al. 2002). From a theoretical perspective, any behavioral effects of racial profiling may arise through different cognitive mechanisms. For example, profiling may increase stereotype-based response tendencies, deteriorate target object discrimination, or both. Identification of the cognitive mechanisms affected by profiling is therefore essential for understanding the origins of weapon identification bias and for evaluating interventions aiming at reductions of racial bias in decision making.

Accordingly, the present research pursues two complementary goals. First, we seek to advance the understanding of the cognitive mechanisms driving racial bias in the WIT. Second, building on this foundation, we investigate whether (and if so, how) these mechanisms are influenced by racial profiling. Together, this should not only advance our theoretical understanding of the WIT but also provide potential leverage for practical interventions and training.

1.1 | Mechanisms of Racially Biased Responding

Although weapon identification bias has been observed reliably (Rivers 2017), the mechanisms underlying this bias have been a matter of debate (Klauer and Voss 2008; Klauer et al. 2015; Klein and Todd 2024; Payne et al. 2005; Todd et al. 2021). One plausible mechanism frequently discussed in the literature is racially biased object discrimination (briefly: discrimination bias) resulting from prime-influenced information extraction during perceptual construal of the target object. The second major candidate is a racially biased response preference (briefly: response bias) induced by prime-triggered threat stereotypes that create an additional stream of information (next to perceptual processing) which varies between primes (Klauer et al. 2015; Klauer and Voss 2008; Payne et al. 2005).

If discrimination bias drives racial bias in the WIT, then prime race should affect the perception and interpretation of parts of the target objects (discrimination bias hypothesis). Specifically, when participants try hastily to resolve perceptual ambiguity, the prime race provides context cues that might foster misperception of the target object. For instance, a metal tube of a target object might be interpreted as the barrel of a gun following a Black prime and as the shaft of a screwdriver following a White prime (Klauer et al. 2015; Klauer and Voss 2008; Payne et al. 2005). This effect should not necessarily persist when participants are given more time for responding. Supporting this account, Payne et al. (2005) asked participants for a second object classification judgement after providing their original response. Notably, in trials with incorrect initial object classifications, they were able to report the correct object when given sufficient time afterwards. Payne and collaborators concluded that target discrimination is not influenced by race, at least if participants are given sufficient time for responding. Additional model-based studies reported mixed results. While some authors found evidence for racially biased object perceptions (Klauer et al. 2015), others did not (Klein and Todd 2024; Todd et al. 2021).

In contrast, if racial bias is mediated by response bias, racial stereotypes elicited by the prime create an additional stream of information from the outset of each trial, that is, before the actual perceptual processing of target objects takes place (response bias hypothesis). Response bias results in default responses triggered by racial stereotypes, that is, a stronger tendency to respond 'gun' after seeing a Black face and to respond 'tool' following a White face. This aligns well with recent research suggesting that prime race influences response bias from the outset of each trial (Klauer and Voss 2008; Laukenmann et al. 2023; Todd et al.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2021). Importantly, because response bias is based on a secondary stream of information, it only determines the factual response if object discrimination fails within the time window provided for responding.

Crucially, these two mechanisms are not mutually exclusive, meaning that a given response pattern may reflect discrimination bias, response bias, or both mechanisms acting in parallel. Furthermore, the relative contribution of each mechanism may change across contexts. Consistent with this, although weapon identification bias can be considered a repeatedly observed, robust effect (Rivers 2017), its strength varies systematically with the conditions under which the task is performed. One important factor is the attention devoted to racial information (i.e., race salience) (Higgins 1996). Manipulations that increase the salience of racial categories strengthen racial bias, whereas those that highlight other features, such as age or emotional expression, attenuate it (Klein and Todd 2024; Klein et al. 2026; Todd et al. 2021). Likewise, explicit instructions to engage in or to avoid racial profiling modulate the bias accordingly (Payne et al. 2002). Determining which underlying mechanism(s) vary depending on racial profiling instructions is thus a key step toward understanding how racial bias can be amplified or contained. One potential way to investigate the underlying mechanisms and their modulation is cognitive modeling via multinomial processing trees (MPTs), as detailed in the following section.

1.2 | Process Dissociation Procedure (PDP) and MPT Modeling

The PDP is a widely used measurement model to disentangle the contributions of discrimination and response biases in WIT performance (Bishara and Payne 2009; Huntsinger et al. 2009; Ito et al. 2015; Klauer et al. 2015; Payne 2001; Payne 2005). Discrimination reflects a participant's latent ability to identify the target object correctly with probability C, whereas response tendency reflects a participant's latent preference to respond *gun* rather than *tool* with probability A (Klauer and Voss 2008; Laukenmann et al. 2023; Payne 2001).

Figure 1 illustrates the standard version of the PDP in the form of two processing trees, one for each target condition (i.e., gun and tool). The left side indicates the to-be-identified target object, followed by the possible processing branches (and their associated probabilities) leading to certain responses in the WIT, as indicated on the right side of the trees. For the gun target tree, for example, discrimination succeeds with probability C, resulting in the correct response *gun*. When discrimination fails with probability $(1 - C)$, response tendencies drive the response *gun* with probability A and the response *tool* with the complementary probability $(1 - A)$. For the tool target tree, the processing branches are analogous, except that successful discrimination results in the correct response *tool*.

Parameters of process dissociation models can be estimated using MPT modeling (for a tutorial, see Schmidt et al. 2025). This approach enables conjoint estimation of the relevant parameters C and A including interval estimates, along with the assessment of model fit and the calculation of information criteria for purposes of selecting among several candidate models. This approach has previously been used to estimate PDP parameters for the WIT (Bishara and Payne 2009; Klauer et al. 2015; Laukenmann et al. 2023).

Note that the PDP historically evolved from the dual-process research tradition where its parameters were conceived as measure of 'controlled' (C) and 'automatic' (A) cognitive processes (e.g., Buchner et al. 1995; Jacoby 1991; Payne 2001). However, we refrain from subscribing to the notion of exactly two distinct types of processes explaining WIT performance. We rather use the PDP as a measurement model to disentangle effects of 'successful discrimination' (C) and 'response tendency toward *gun*' (A) mechanisms driving WIT responses. Conceived in this way, the PDP framework remains agnostic regarding the number and psychological nature of cognitive processes involved in WIT performance (Klauer and Voss 2008; Laukenmann et al. 2023). Hence, it is compatible with different process theories proposed for sequential priming tasks, including single-process (De Houwer 2014; Klauer and Voss 2008; Klein and Todd 2024; Todd et al. 2021), dual-process (Gawronski and Bodenhausen 2006; Payne 2001; Klauer and Voss 2008; Laukenmann et al. 2023) and even multi-process accounts (Conrey et al. 2005).

1.3 | Hypothesis Testing and Auxiliary Assumptions

How can perceptual discrimination and response preferences mediate racial bias in the WIT? If racially biased response preferences are at work this should cause a higher rate of *gun* responses following Black face primes, given discrimination failure. This will be mirrored in a larger A-parameter for the Black (B) than for the White face prime (W) condition, hence $A_B > A_W$ (Klauer et al. 2015; Ito et al. 2015; Payne 2001). In contrast, if racially biased object discrimination is involved, then prime race affects object discrimination. Hence, parameter C is expected to vary as a function of both prime and target conditions. In the Black prime condition, the prime is predicted to enhance target discrimination for guns (G) compared to tools (T), resulting in a larger C-parameter for gun than for tool targets, that is, $C_{BG} > C_{BT}$. In the White prime condition, in contrast, the prime is predicted to have opposite effects, resulting in $C_{WG} < C_{WT}$ (Klauer and Voss 2008; Klauer et al. 2015). Both order restrictions can be formalized by means of shrinkage parameters in the model equations (Knapp and Batchelder 2004). For example, in the Black prime condition, the order restriction $C_{BG} \geq C_{BT}$ is implemented by the constraint $C_{BT} = s_B \cdot C_{BG}$, with $0 \leq s_B \leq 1$, serving as shrinkage parameter.

Note that the standard version of the PDP illustrated in Figure 1 cannot accommodate this prime by target interaction in the C-parameter because successful discrimination is assumed to be invariant across target conditions (implying $C_{BG} = C_{BT} = C_B$ and $C_{WG} = C_{WT} = C_W$). Hence, to evaluate whether discrimination bias, response bias, or both are responsible for racially biased responding in the WIT, a generalization of the standard PDP model is necessary that allows the C-parameter to vary between prime and target conditions. However, developing such a generalized PDP variant from the model in Figure 1 would result in an overparameterized model when applied to two types of primes crossed with two target conditions: It would include six parameters (i.e., four C-parameters and two A-parameters)

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

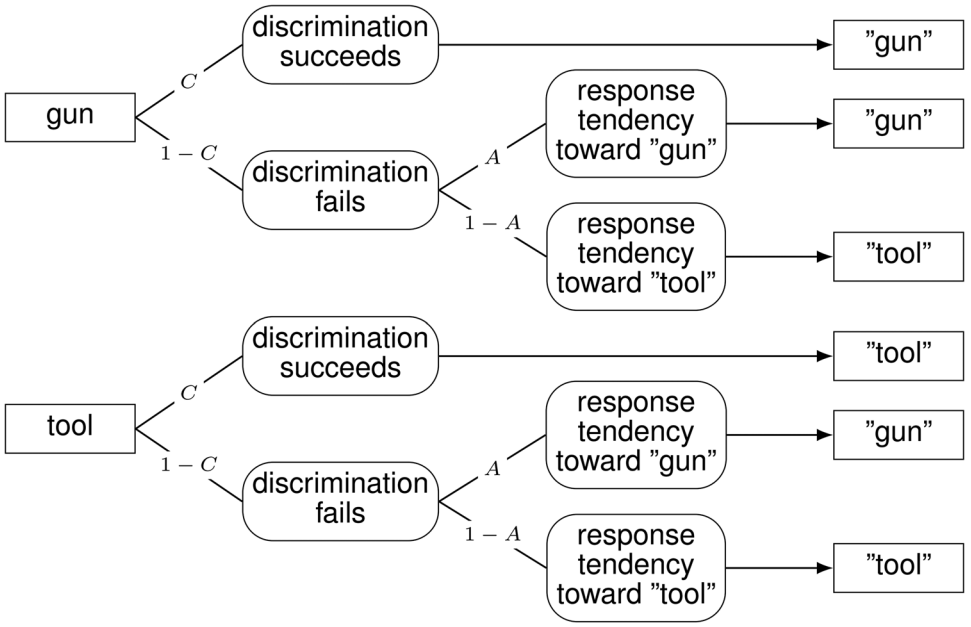

FIGURE 1 | Standard PDP model for the Weapon Identification Task. Abbreviation: PDP, process dissociation procedure. Figure adapted from Laukenmann et al. (2023). Parameters *C* and *A* denote probabilities of successful object discrimination and a response tendency toward 'gun', respectively. Note that *A* is conditional on discrimination failure, that is, *A* represents the conditional probability of a 'gun' response tendency given discrimination fails. Also note that the standard PDP model illustrated here is based on the invariance assumption, that is, both *C* and *A* are assumed to be invariant across target objects (Klauer et al. 2015).

but provide only four independent category probabilities (i.e., probabilities of correct responses for each of the four prime-target combinations under investigation). Obviously, such a model would be technically nonidentifiable.

More degrees of freedom for PDP modeling can be gained by including an additional within-subject manipulation that leaves the parameter *C* unaffected. Klauer et al. (2015) previously used this approach, based on the assumption that pay-off manipulations for target objects (i.e., different rewards for correct *gun* and correct *tool* responses) affect the overall response tendency *A* selectively (i.e., only the general tendency to respond *gun* vs. *tool*). Their results supported the discrimination bias hypothesis, that is, *C* varied in their study not only between primes but also between targets in the manner described above. To check the generalizability of their results, we manipulated the general (i.e., prime-independent) response tendency selectively by employing different base-rates of gun and tool targets in the WIT to achieve an identifiable generalized PDP model. The same manipulation was previously used successfully in other applications of the PDP (Buchner et al. 1995; Klauer et al. 2003) and similar procedures such as the Implicit Association Test (Conrey et al. 2005). Like Klauer et al.'s (2015) pay-off manipulation, increasing the proportion of gun (or tool) targets in the WIT will certainly affect the general response tendency in favor of *gun* (or *tool*) responses as reflected in the *A*-parameter of the PDP, but likely not discrimination bias reflected in the *C*-parameter. We return to this assumption in the General Discussion section.

Based on an extended WIT data structure that encompasses low and high base-rates of gun targets, we can thus go beyond past research and investigate whether (1) the racial bias effect is reflected in the prime-target interaction for the *C*-parameter as predicted by the discrimination bias hypothesis, (2) in the prime main effect for the *A*-parameter as predicted by the response bias hypothesis, or (3) both.[1]

2 | Research Overview

We conducted three pre-registered studies to investigate discrimination and response biases as potential mechanisms underlying racial bias in the WIT (Klauer et al. 2015; Payne et al. 2005). Study 1 used standard WIT instructions and made use of a base-rate manipulation for target objects to allow application of the generalized PDP model. Study 2 replicated and extended Study 1. In addition, Study 2 investigated whether the instruction to engage in racial profiling to facilitate target identification ('Use Race' instruction, see Payne et al. 2002) moderates the processes involved in the racial bias effect. Study 3 replicated and extended Studies 1 and 2. Specifically, Study 3 investigated whether the instruction to engage in racial profiling ('Use Race' instruction) versus the instruction to avoid racial profiling ('Avoid Race' instruction, see Payne et al. 2002) differently affect the processes underlying racial bias effects. For all three studies, we report how we determined our sample size, all data exclusions, all manipulations and all measures. The pre-registrations, data, analytic scripts and further analyses are available in the Open Science Framework (OSF, see https://osf.io/4yxv9). In line with previous research (Klein and Todd 2024; Payne et al. 2005; Todd et al. 2021), we predicted that response bias drives weapon identification bias in the WIT. Since Studies 1 and 2 confirmed this prediction for standard WIT instructions but not for racial profiling instructions (where discrimination bias was involved in addition to response bias, see below), we performed Study 3 to replicate the patterns observed in Studies 1 and 2 and

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to additionally explore effects of an instruction that aims at preventing racial profiling (the avoid race instruction).

3 | Study 1

Study 1 consisted of two within-subject conditions of the WIT with different base-rates of gun and tool targets. First, we investigated whether weapon identification bias can be observed in error rates (Payne 2001; Rivers 2017). Second, we used a generalized PDP model to determine which mechanism can best account for the *C*- and *A*-parameter differences across prime and target conditions.

3.1 | Method

3.1.1 | Participants

We aimed for a sample size of 170 participants to match the sample size used in Study 5 of Klauer et al. (2015) who performed a similar model comparison. To achieve this sample size and control for participants' performance in an online study, we continued data collection until our target sample size was reached after excluding participants based on the pre-registered exclusion criteria. Specifically, participants were excluded when they failed the attention check and when their percentage of missing responses was an extreme outlier according to Tukey's criterion (i.e., 1.5 times the interquartile range above the median, see Clark-Carter 2004, Chapter 9; Klauer et al. 2015). We collected data of 211 participants (missing responses: mean = 14.2%, SD = 26.0%, median = 2.0%). We excluded 41 participants with more than 25.8% missing responses or failed attention check, resulting in a final sample size of $N = 170$. Mean age was 24.1 years (SD = 7.0), and participants indicated their gender as 129 female, 34 male and 7 other. The self-reported race of participants was 7.6% Black, 61.2% White and 31.2% other.

3.1.2 | Design

Independent variables were base-rate (more guns [mg] vs. more tools [mt]), prime face (Black vs. White) and target object (gun vs. tool). All three independent variables were completely cross-classified in a balanced 2 × 2 × 2 within-subject design. The participants' error rates per condition (proportion of incorrect '*gun*' or '*tool*' responses, depending on trial type) served as the dependent variable.

3.1.3 | Materials

Prime faces, target objects and pattern mask were taken from Rivers (2017). Targets were five drawings of hand tools and five drawings of weapons presented in four different orientations rotated by 90°. Target objects and the pattern mask can be accessed via OSF at: https://osf.io/9e6sa/. Prime faces consisted of 24 Black and 24 White male faces with neutral facial expression, plus one neutral prime consisting of the outline of a face which was only used for the practice trials. Prime faces were provided by the first author of Phills et al. (2011). Each image was displayed in 300 × 300-pixel format.

3.1.4 | Procedure

Data collection was conducted online via Prolific (www.prolific.co) (2021). Participants were required to be at least 18 years of age and be part of the Prolific US-sample. The study took about 25 min and participants were rewarded with £3.20. After providing consent, participants completed the WIT.

The WIT consisted of 24 practice trials, follow by 200 blocked experimental tirals for each base-rate condition (i.e., with prime-target combinations balanced for race). Trials were assinged to each participant in random order within base-rate conditions. The mt condition consisted of a 70%:30% ratio of tools relative to guns. Correspondingly, the mg condition consisted of a 30%:70% ratio. Participants were informed about the percentage of tool and gun targets beforehand. As attention check, participants were instructed to type the percentage of targets they will see in an open text field on the following screen. To rule out order effects, the base-rate conditions were presented in random order for each participant.

In each trial, participants saw a sequence of a fixation cross (500 ms), a face prime (200 ms), a target object (200 ms), a pattern mask (300 ms) and a feedback screen (1000 ms) in the centre of the screen. In practice trials, possible feedback was 'correct', 'false' or 'too slow'. In experimental trials, either no feedback occurred or 'too slow' was provided as feedback. Participants had a response time limit of 500 ms after target onset. Participants were instructed to identify as fast and accurately as possible the target object while ignoring the face prime. For each participant, responding with 'gun' or 'tool' was randomly assigned to response keys *D* and *L*. After completing the WIT, participants completed a basic demographics questionnaire (age, gender, and race) and were asked what they thought the study was about. At the end of the study, participants were thanked and debriefed.

3.1.5 | MPT Modeling Procedure

We used the Bayesian hierarchical latent-trait MPT approach of Klauer (2010) as implemented in TreeBUGS (Heck et al. 2018) to estimate the PDP models.[2] For assessing goodness-of-fit, we used the test statistics T_1 and T_2 to calculate Bayesian posterior predictive *p*-values. T_1 summarizes how well the model accounts for the average response frequencies across participants. T_2 summarizes how well the model accounts for the variances and correlations of the response frequencies across participants. The posterior predictive *p*-value of either of these statistics represents the comparison of the calculated test statistics obtained for observed and predicted response frequencies, with a value of $p > 0.05$ typically seen as evidence that model assumptions are in line with the data (Klauer 2010; Klauer et al. 2015).

For Bayesian model comparisons, we used the Deviance Information Criterion (DIC) (Klauer et al. 2015; Spiegelhalter et al. 2002) and the Widely Applicable Information Criterion (WAIC) (Vehtari et al. 2017; Watanabe 2010) as information criteria taking model complexity into account. The model with the lowest DIC (WAIC) value represents the best compromise between model fit

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Proportion of errors by prime, target and base-rate condition in Study 1.

	Prime			
	Black		White	
Target	***M***	***SD***	***M***	***SD***
More guns condition				
Gun	0.161	0.103	0.181	0.102
Tool	0.544	0.202	0.511	0.185
More tools condition				
Gun	0.502	0.197	0.518	0.184
Tool	0.165	0.098	0.146	0.096

Abbreviation: *M* = Mean; *SD* = standard deviation.

TABLE 2 | Goodness-of fit statistics and model selection measures for Study 1.

Parameter constraints	$p(T_1)$	$p(T_2)$	DIC	WAIC
Unconstrained	0.49	0.08	7215.0	7035.1
$A_{Bmg} = A_{Wmg}; A_{Bmt} = A_{Wmt}$	0.06	< 0.001	7385.9	7360.0
$C_{BG} = C_{BT} = C_{WG} = C_{WT}$	< 0.001	< 0.001	7405.1	7386.4
$C_{BT} = s_{B1} \cdot C_{BG}; C_{WT} = s_{W1} \cdot C_{WG}$	0.14	0.01	7276.3	7144.9
$s_{B1} = s_{W1}$	0.13	0.01	7262.5	7136.7
$s_{B1} \leq s_{W1}$	0.11	0.01	7266.0	7140.4

Abbreviations: *A* = response tendency parameter, *B* = Black prime, *C* = successful discrimination parameter, *G* = gun target, mg = more guns target condition, mt = more tools target condition, *s* = shrinkage parameter, *T* = tool target, *W* = White prime.

and model complexity. A ΔDIC (ΔWAIC) difference larger than two is typically interpreted as evidence for one model compared to the other (Burnham and Anderson 2004; Klauer et al. 2015; Spiegelhalter et al. 2002). We used the R package TreeBUGS (Heck et al. 2018). The Markov Chain Monte Carlo (MCMC) algorithm was run for three independent estimation chains with 1,000,000 iterations each of which 100,000 were removed as a burn-in. Every 100th iteration was retained to compute summary statistics. The Rubin–Gelman statistics $\hat{R}$ was smaller than 1.05 for all parameter estimates across all models, showing acceptable convergence.

3.2 | Results

Prior to all analyses, we excluded trials with latencies < 100 ms and > 1500 ms, resulting in exclusion of 3.8% of the trials. Given $N = 170$ participants, a sensitivity power analysis revealed that a relatively small within-subject population effect size of $\eta_p^2 = 0.045$ (see Cohen 1988)[3] can be detected in repeated-measures analysis of variance (ANOVA) $F(1, 169)$ tests with a Type-1 error level of $\alpha = 0.05$ and a statistical power of $1 - \beta = 0.80$ (Faul et al. 2009). This holds for any main or interaction $F(1, 169)$ test (and simple main effect *t*-test) in Study 1. Thus, our study is sufficiently powered.

3.2.1 | Error Rates

Table 1 shows that for both base-rate conditions, mean error rates for tool targets were higher for Black compared to White primes while the opposite holds for gun targets. As expected, a 2 (prime) × 2 (target) × 2 (base-rate) repeated-measures ANOVA reveals a significant prime × target interaction, $F(1, 169) = 10.90$, $p = 0.001$, $\eta_p^2 = 0.061$, in line with the pattern indicative of racial bias. We report all further significant and non-significant interactions in the Appendix 2.

3.2.2 | PDP Modeling

Table 2 summarizes the model fit statistics, that is, the posterior predictive *p*-values (i.e., $p(T_1)$, $p(T_2)$) and the model selection measures taking model complexity into account (i.e., DIC, WAIC) for different versions of the generalized PDP model. The unconstrained PDP model with four *C*-parameters (varying between the four prime-target conditions but equated across base-rates) and four *A*-parameters (varying between all prime-base-rate combinations but equated across target items) fits the data well and shows best model selection measures. We therefore use this model as our base model and report parameter estimates accordingly.

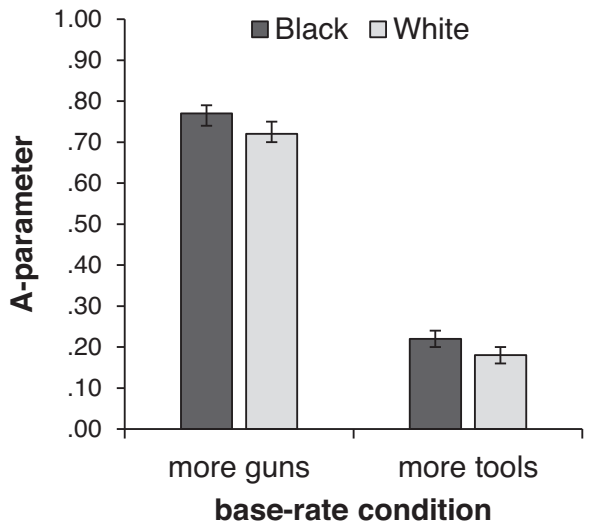

FIGURE 2 | *A*-parameter estimates of Study 1 by prime race and base-rate condition. *Note:* Mean parameter estimates with 95% Bayesian credible intervals (BCIs) as error bars.

Response bias: As illustrated in Figure 2, mean *A*-parameter estimates for Black primes were descriptively higher compared to White primes in both the mg ($A_{Bmg} = 0.77$ [0.74–0.79]; $A_{Wmg} = 0.72$ [0.70–0.75])[4] and the mt condition ($A_{Bmt} = 0.22$ [0.20–0.24]; $A_{Wmt} = 0.18$ [0.16–0.20]). To test whether *A*-parameters for Black primes are credibly larger than for White primes, we estimated a constrained model which equates *A*-parameters across prime conditions within each base-rate condition (i.e., $A_{Bmg} = A_{Wmg}$ and $A_{Bmt} = A_{Wmt}$). Model comparison shows better fit of the unconstrained model compared to the constrained model, indicating an influence of race on response bias as predicted (see Table 2).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

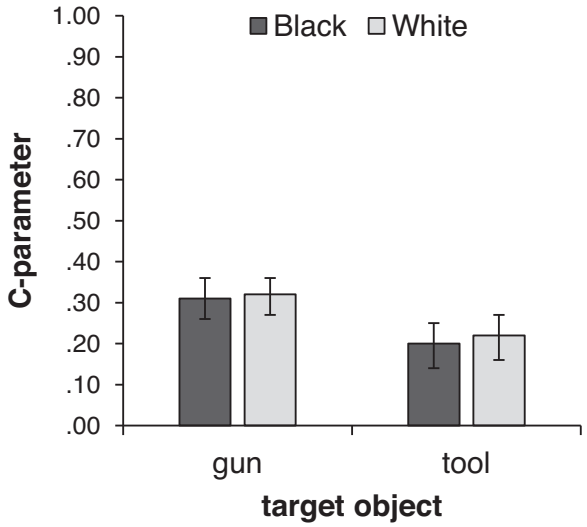

FIGURE 3 | *C*-parameter estimates of Study 1 by prime race and target object. *Note:* Mean parameter estimates with 95% BCIs as error bars.

Discrimination bias: As illustrated in Figure 3, mean *C*-parameter estimates for gun targets were descriptively higher compared to tool targets in both the Black prime condition (C_{BG} = 0.31 [0.26–0.36]; C_{BT} = 0.20 [0.14–0.25]) and the White prime condition (C_{WG} = 0.32 [0.27–0.36]; C_{WT} = 0.22 [0.16–0.27]). While the former result is in line with the discrimination bias hypothesis, the latter result is not. To test whether a relative discrimination advantage of guns compared to tools in the Black prime condition is larger than in the White prime condition (i.e., whether there is at least an ordinal interaction between primes and targets on *C*-parameters), we compared model fit of a baseline model assuming $C_{BT} = s_{Bl} \cdot C_{BG}$ and $C_{WT} = s_{Wl} \cdot C_{WG}$ (reflecting an ordinal interaction) to a model with the same parameter structure and the additional equality constraint $s_{Bl} = s_{Wl}$ (reflecting a main effect of target object; see Kuhlmann et al. 2019). The comparison revealed that the model with the equality constraint ($s_{Bl} = s_{Wl}$) fits the data better than both the model without this constraint (s_{Bl}; s_{Wl}) and the ordinal prime-target interaction model ($s_{Bl} \leq s_{Wl}$). Thus, there is no support for the discrimination bias hypothesis according to which target discrimination should be differently affected by prime race. Rather, our results support the idea of a generally higher probability of successful discrimination for gun compared to tool targets, irrespective of prime. In line with this, a model constraining all *C*-parameters to be equal (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$) does not fit the data well (see Table 2).

3.3 | Discussion

Results of Study 1 supported the hypothesis that racial bias in the WIT is mediated by response bias while they were inconsistent with the discrimination bias hypothesis. To reiterate, these results rest on the assumption that the base-rate manipulation of target objects selectively affects response tendency and leaves discrimination bias unaffected. Notably, this is a standard assumption often made in recognition and discrimination research (e.g., Buchner et al. 1995; Conrey et al. 2005).

Although Study 1 clearly shows that response bias is the driving force that triggers weapon identification bias, it is silent about the generalizability of this finding. For example, it remains an open question whether the finding still holds when participants are prompted to use racial profiling as a strategy to identify the target object. Previous work has shown that racial profiling strengthens weapon identification bias (Payne et al. 2002), but did not disentangle its effects on underlying mechanisms such as discrimination bias and response bias. We therefore investigated this research question in Study 2.

4 | Study 2

Study 2 included two experimental groups that differed in task instructions. The control group of Study 2 was a direct replication of Study 1. The 'Use Race' group received an additional instruction to engage in racial profiling for target discrimination (see Appendix A1). Hence, Study 2 allowed us to assess the replicability of Study 1 on the one hand and the generalizability of results to participants working under 'Use Race' instructions on the other hand.

4.1 | Method

4.1.1 | Participants

As in Study 1, we aimed at a sample size of 170 participants for each of the two experimental groups. We used the same data collection strategy as in Study 1, including an attention check test for the use race group. We collected data of 428 participants (missing responses for the control group: mean = 7.2%, SD = 13.2%, median = 1.8%; for the use race group: mean = 7.9%, SD = 15.4%, median = 1.3%). We excluded 94 participants with missing responses above Tukey's criterion or failed attention check (control: > 17.6% missing responses, N = 37; use race: > 15.1% missing responses, N = 57), resulting in N = 169 for the control and N = 165 for the use race group.[5] Mean age was 33.8 years (SD = 12.4) in the control, and 36.5 years (SD = 12.1) in the use race group. Participants indicated their gender as 88 (78) female, 76 (84) male and 5 (3) other, in the control (use race) group, respectively. Self-reported race was 5.9% (12.7%) Black, 66.3% (69.7%) White and 27.8% (17.6%) other in the control (use race) group, respectively.

4.1.2 | Design

Independent variables were instruction group (control vs. use race, between-subjects), base-rate (mg vs. mt, within-subject), prime face (Black vs. White, within-subject) and target object (gun vs. tool, within-subject). All four independent variables were completely cross-classified in a balanced 2 × 2 × 2 × 2 mixed design. Error rates per condition served as the dependent variable.

4.1.3 | Materials

Target images were the same as in Study 1. The use race instruction enforcing racial profiling (see Appendix A1) was taken from Payne et al. (2002).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.1.4 | Procedure

Data collection was conducted online via Prolific with the same sample specifications as in Study 1. The study took about 28 min and participants were rewarded £3.50. The study procedure was the same as in Study 1 except that the use race group received the additional racial profiling instruction after they provided their consent. As additional attention check test in this condition, participants had to indicate the experimental group they have been assigned to in an open text field presented on the screen that followed the use race instruction.

4.1.5 | MPT Modeling Procedure

The modeling procedure was the same as in Study 1. Models were estimated separately for the control and the use race group. Again, the Rubin–Gelman statistics $\hat{R}$ was smaller than 1.05 for all parameter estimates and all models, showing an acceptable convergence of MCMC sampling.

4.2 | Results

Prior to all analyses, we excluded trials with latencies < 100 ms and > 1500 ms resulting in the exclusion of 2.5% of the trials. Given $N = 165$ in the use race group, a sensitivity power analysis revealed that a relative small within-subject population effect size of $\eta_p^2 = 0.046$ (see Cohen 1988) can be detected in repeated measures ANOVA $F(1, 164)$ tests with a Type-1 error level of $\alpha = 0.05$ and a statistical power of $1 - \beta = 0.80$ (Faul et al. 2009). This holds for any main or interaction $F(1, 164)$ test (and simple main effect t-test) for the use race group. For the control group with $N = 169$, power is similar. For the $F(1, 332)$ tests involving both experimental groups, statistical power is even higher. Thus, like Study 1, Study 2 is well powered.

4.2.1 | Error Rates

Table 3 shows that the racial bias pattern in observed error rates replicates again and is descriptively stronger in the racial profiling group. A 2 (instruction) × 2 (prime) × 2 (target) × 2 (base-rate) mixed-design repeated-measures ANOVA revealed a significant instruction × prime × target triple interaction, $F(1, 332) = 22.6$, $p < 0.001$, $\eta_p^2 = 0.064$. To decompose this interaction, we conducted simple-effects analyses for each instruction group. The prime × target interaction was significant for both the control group, $F(1, 332) = 10.53$, $p = 0.001$, $\eta_p^2 = 0.031$, and the use race group, $F(1, 332) = 97.81$, $p < 0.001$, $\eta_p^2 = 0.228$. Hence, the significant triple interaction reflects a stronger racial bias effect in the use race group, indicating more pronounced weapon identification bias due to racial profiling instructions. For the complete ANOVA tables, see the Appendix 2.

TABLE 3 | Proportion of errors by prime, target and base-rate condition in Study 2.

	Prime			
	Black		White	
Target	***M***	**SD**	***M***	**SD**
Control group				
More guns condition				
Gun	0.138	0.087	0.150	0.091
Tool	0.502	0.204	0.483	0.203
More tools condition				
Gun	0.440	0.192	0.481	0.206
Tool	0.156	0.098	0.125	0.088
Use race group				
More guns condition				
Gun	0.128	0.102	0.188	0.118
Tool	0.529	0.219	0.447	0.215
More tools condition				
Gun	0.405	0.209	0.520	0.186
Tool	0.197	0.136	0.136	0.108

Abbreviation: *M* = Mean; *SD* = standard deviation.

4.3 | PDP Modeling

The unconstrained generalized PDP model fitted the data well for both the control and the use race group. It also showed the best model selection indices (see Table 4). We therefore used this model as our baseline model and report parameter estimates accordingly. Again, as in Study 1, to evaluate racial bias effects on A- and C-parameters, we compared model fit and model selection measures between the baseline model and restricted models equating A- or C-parameters indicative of racial bias (see Table 4).

Response bias: As illustrated in Figure 4, in the control group, mean A-parameter estimates for Black primes were descriptively higher compared to White primes for both the mg ($A_{Bmg} = 0.77$ [0.75–0.79]; $A_{Wmg} = 0.75$ [0.72–0.77]) and the mt ($A_{Bmt} = 0.23$ [0.21–0.25]; $A_{Wmt} = 0.18$ [0.16–0.19]) condition. As in Study 1, model comparison shows better fit of the unconstrained model compared to the constrained model that equates A-parameters across prime conditions within base-rate conditions, indicating an influence of race on response bias.

In the use race group, mean A-parameter estimates for Black primes were also descriptively higher compared to White primes for both the mg ($A_{Bmg} = 0.78$ [0.75–0.81]; $A_{Wmg} = 0.73$ [0.71–0.76]) and the mt ($A_{Bmt} = 0.26$ [0.24–0.29]; $A_{Wmt} = 0.20$ [0.18–0.22]) condition. Model comparison again shows a better fit for the unconstrained model compared to the constrained model that equates A-parameters across prime conditions within base-rate conditions, reflecting an influence of race on response bias in the use race group.

Discrimination bias: As illustrated in Figure 5, in the control group, mean C-parameter estimates for gun targets were descriptively higher compared to tool targets in both the Black prime condition ($C_{BG} = 0.39$ [0.34–0.44]; $C_{BT} = 0.28$ [0.23–0.34]) and the White prime condition ($C_{WG} = 0.38$ [0.33–0.43]; $C_{WT} = 0.29$ [0.23–0.34]). Again, to test for a prime-target interaction on C-

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Goodness-of fit statistics and model selection indices for Study 2.

Parameter constraints	Control group				Use race group			
	$p(T_1)$	$p(T_2)$	DIC	WAIC	$p(T_1)$	$p(T_2)$	DIC	WAIC
Unconstrained	0.47	0.13	7049.4	6902.8	0.51	0.26	6952.4	6731.8
$A_{Bmg} = A_{Wmg}; A_{Bmt} = A_{Wmt}$	< 0.001	0.08	7069.9	6989.7	< 0.001	< 0.001	7349.8	7397.2
$C_{BG} = C_{BT} = C_{WG} = C_{WT}$	< 0.001	0.01	7202.3	7195.9	< 0.001	0.01	7055.9	6957.0
$C_{BT} = s_{BI} \cdot C_{BG}; C_{WT} = s_{WI} \cdot C_{WG}$	0.21	0.03	7094.6	6695.4				
$s_{BI} = s_{WI}$	0.24	0.03	7083.6	6987.9				
$s_{BI} \leq s_{WI}$	0.21	0.03	7084.7	6990.9				
$C_{BT} = s_{BI} \cdot C_{BG}; C_{WG} = s_{W2} \cdot C_{WT}$					0.42	0.09	6984.2	6812.1
$C_{BT} = s_{BI} \cdot C_{BG}; C_{WG} = C_{WT}$					0.42	0.10	6984.3	6812.3
$C_{BT} = C_{BG}; C_{WG} = s_{W2} \cdot C_{WT}$					0.42	0.09	6983.6	6812.6

Abbreviations: A = response tendency parameter, B = Black prime, C = successful discrimination parameter, G = gun target, mg = more guns target condition, mt = more tools target condition, s = shrinkage parameter, T = tool target, W = White prime.

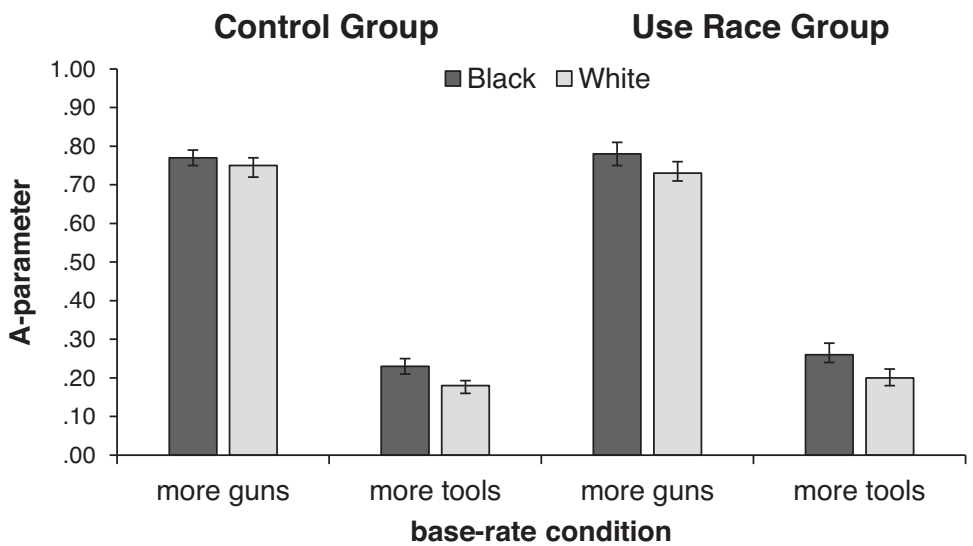

FIGURE 4 | *A*-parameter estimates of Study 2 by prime race, base-rate condition and instruction condition. *Note:* Mean parameter estimates with 95% BCIs as error bars.

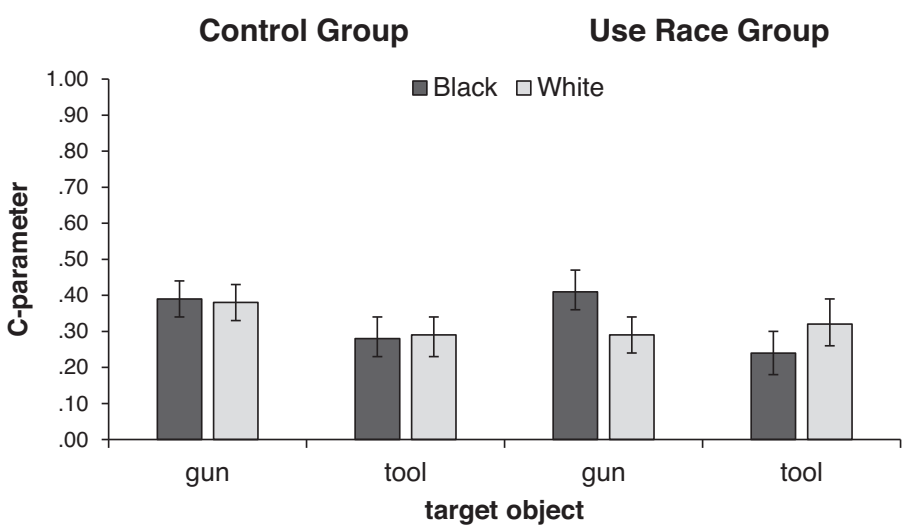

FIGURE 5 | *C*-parameter estimates of Study 2 by prime race, target object and instruction condition. *Note:* Mean parameter estimates with 95% BCIs as error bars.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

parameters, we compared model fit of a baseline model assuming $C_{BT} = s_{BI} \cdot C_{BG}$ and $C_{WT} = s_{WI} \cdot C_{WG}$ (reflecting an ordinal interaction) to a model with the same parameter structure and the additional equality constraint $s_{BI} = s_{WI}$ (reflecting a main effect of target object). The comparison revealed that, as in Study 1, the model with the equality constraint ($s_{BI} = s_{WI}$) fitted the data better than both the model without this constraint (s_{BI}; s_{WI}) and the ordinal prime-target interaction model ($s_{BI} \leq s_{WI}$). In the control group of Study 2, there is thus no support for the discrimination bias hypothesis according to which racial bias is mediated by object discrimination. Replicating Study 1, the data suggest a generally higher probability of successful discriminations for gun compared to tool targets, irrespective of prime. In line with this, a model constraining all *C*-parameters to be equal (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$) did not fit the data well.

In the use race group, mean *C*-parameter estimates for gun targets were descriptively higher compared to tool targets in the Black prime condition (C_{BG} = 0.41 [0.36–0.47]; C_{BT} = 0.24 [0.18–0.30]) but reversed in the White prime condition (C_{WG} = 0.29 [0.24–0.34]; C_{WT} = 0.32 [0.26–0.39]). Hence, these estimates display the disordinal interaction expected under the discrimination bias hypothesis (i.e., $C_{BG} > C_{BT}$ and $C_{WG} < C_{WT}$). To test this interaction, we first compared a model assuming $C_{BT} = s_{BI} \cdot C_{BG}$ and $C_{WG} = s_{W2} \cdot C_{WT}$ (reflecting a disordinal interaction) to a model assuming all *C*-parameters to be equal (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$). The comparison revealed that the disordinal model fitted the data better than the model assuming equal *C*-parameters. For a more rigorous assessment of the disordinal interaction hypothesis, we additionally estimated a model assuming a discrimination advantage restricted to guns in the Black prime condition (i.e., $C_{BT} = s_{BI} \cdot C_{BG}$ and $C_{WG} = C_{WT}$) and a model assuming a discrimination advantage restricted to tools in the White prime condition (i.e., $C_{WG} = s_{W2} \cdot C_{WT}$ and $C_{BT} = C_{BG}$). Neither the former nor the latter model provides evidence of better model fit compared to the disordinal interaction model. Thus, the use race instruction biases target discrimination in both prime conditions. Overall, these results clearly support the discrimination bias hypothesis in the use race instruction group.

4.4 | Discussion

Replicating Payne et al. (2002), a weapon identification bias modulated by race was observed across both base-rate conditions for both instruction groups, with a more pronounced weapon identification bias when participants engage in racial profiling. The modeling results for the control group successfully replicated Study 1 and supported the hypothesis that racially biased WIT performance is driven by response bias rather than by discrimination bias. In contrast, the results for the use race group showed that both response bias and discrimination bias are involved when a WIT participant is prompted to engage in racial profiling.

Although the control group of Study 2 successfully replicates Study 1, the double-bias pattern observed in the use race instruction group of Study 2 (i.e., response bias plus discrimination bias underlying the WIT) remains to be replicated. This is particularly important because, due to the strict response deadline of only 500 ms, error rates were relatively high in both Study 1 and Study 2. It would be interesting to see whether the result pattern generalizes to more liberal response deadlines associated with lower error rates. Moreover, Study 2 does not reveal which aspect of the rather complex 'Use Race' instruction is responsible for triggering discrimination bias. Is it the mere information about an alleged effect of racial profiling as included in the first part of the instruction (i.e., 'Research has shown that the race of the face sometimes impacts the ways that people classify the second object. People are sometimes faster and more accurate in responding to guns after a Black face than after a White face.')? Or is it the requirement to engage in racial profiling as endorsed in the second part (i.e., 'Regardless of your personal views, we would like you to play the role of someone engaged in racial profiling.')? We designed Study 3 to clarify these questions.

5 | Study 3

Study 3 included three experimental groups that differed in task instructions. The control group was a conceptual replication of Study 1 and the control group in Study 2. Analogously, the 'Use Race' group was a conceptual replication of the corresponding instruction group in Study 2. The 'Avoid Race' group was instructed to avoid (rather than use) racial profiling (see Appendix A1). As in the 'Use Race' group, racial information was salient in the 'Avoid Race' group, because both groups were instructed 'that the race of the face sometimes impacts the ways people classify objects'. However, in contrast to the use race group, the avoid race group was subsequently instructed '(…) to base your responses only on whether the second object looks more like a gun or tool. Try not to let the race of the face influence your decisions to much'. Hence, Study 3 allowed us to assess replicability of previous results and, in addition, to investigate whether informing participants about racial profiling followed by the instruction to use or to avoid race information moderates the mechanisms mediating racial bias in the WIT.

5.1 | Method

5.1.1 | Participants

As in Studies 1 and 2, we aimed at a sample size of 170 participants for each of the three experimental groups. We used the same data collection strategy as in Studies 1 and 2, including an attention check test for the use and avoid race instruction groups. We collected data of 659 participants (missing responses for the control group: mean = 2.6%, SD = 9.5%, median = 0.3%; for the use race group: mean = 2.6%, SD = 8.5%, median = 0.3%; for the avoid race group: mean = 3.8%, SD = 13.2%, median = 0.3%). We excluded 148 participants with missing responses exceeding Tukey's criterion or who failed the attention check (control: > 2.5% missing responses, N = 34; use race: > 3.1% missing responses, N = 48; avoid race: > 2.5% missing responses, N = 66), resulting in a final sample N = 511 (control: N = 170; use race: N = 171; avoid race: N = 170). Mean age was 39.4 years (SD = 13.6), 37.4 years (SD = 12.1) and 39.3 years (SD = 13.0) in the control, the use race and the avoid race groups, respectively. Participants indicated their gender as 93 (93; 94) female, 75 (76; 72) male and 2 (2; 4) other and race as 16.5% (22.8%; 14.7%) Black, 62.4% (49.7%; 61.2%) White and 21.2% (27.5%; 24.1%) other in the control (use race; avoid race) group, respectively.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.1.2 | Design

Independent variables were group (control vs. use race vs. avoid race, between-subjects), base-rate (mg vs. mt, within-subject), prime race (Black vs. White, within-subject) and target object (gun vs. tool, within-subject). All four independent variables were completely cross-classified in a balanced 3 × 2 × 2 × 2 mixed design. The participants' error rates per condition served as the dependent variable.

5.1.3 | Materials

Target images were the same as in Studies 1 and 2. The use race and the avoid race instructions (see Appendix A1) were adopted from Payne et al. (2002).

5.1.4 | Procedure

Data collection was conducted online via Prolific with the same sample specifications as in Studies 1 and 2. The study took about 28 min and participants were rewarded £4.20. The study procedure was the same as in Studies 1 and 2 except for three aspects. First, the use and avoid race groups received the respective additional instruction after they provided their consent. As an additional attention check, participants had to indicate their experimental group in a text field that followed the additional instruction. Second, we changed the base-rate to a 60%:40% ratio of tools relative to guns in the mt condition, and to a 40%:60% ratio in the mg condition. Third we changed the response time limit to 700 ms after target onset by presenting the pattern mask for 500 ms (instead of 300 ms). We changed the base-rate ratios and the response time limit for two reasons: first, to check for the generalizability of findings across different base-rate conditions and response time limits, and second to allow participants to execute more controlled responding resulting in lower overall error rates compared to Studies 1 and 2.

5.1.5 | MPT Modeling Procedure

The modeling procedure was the same as in Studies 1 and 2 except that we increased MCMC sampling parameters to achieve acceptable convergence of MCMC sampling. Resulting in 10,000,000 iterations, with the first 1,000,000 iterations removed as burn-in and every 1000th iteration retained to compute summary statistics. Models were estimated separately for the control, use race and avoid race group. As in Studies 1 and 2, the Rubin–Gelman statistics $\hat{R}$ was smaller than 1.05 for all parameter estimates across all models, showing an acceptable convergence of MCMC sampling.

5.2 | Results

Prior to all analyses, we excluded trials with latencies < 100 ms and > 1700 ms resulting in the exclusion of 0.4% of the trials. Within each of the three experimental groups of participants, the statistical power of the $F(1, 169)$ tests (and simple main effects t-tests) is identical to Study 1. For the $F(2, 508)$ tests including all three experimental groups, statistical power is even higher. Thus, like Studies 1 and 2, Study 3 is well powered. Mean error rates and standard deviations are summarized in Table 5.

TABLE 5 | Proportion of errors by prime, target and base-rate condition in Study 3.

	Prime			
	Black		White	
Target	***M***	**SD**	***M***	**SD**
Control group				
More guns condition				
Gun	0.069	0.089	0.078	0.094
Tool	0.127	0.115	0.104	0.084
More tools condition				
Gun	0.112	0.111	0.130	0.108
Tool	0.068	0.064	0.055	0.055
Use race group				
More guns condition				
Gun	0.065	0.074	0.090	0.097
Tool	0.153	0.156	0.115	0.104
More tools condition				
Gun	0.114	0.095	0.149	0.128
Tool	0.097	0.114	0.064	0.067
Avoid race group				
More guns condition				
Gun	0.067	0.074	0.082	0.079
Tool	0.118	0.111	0.107	0.110
More tools condition				
Gun	0.121	0.107	0.147	0.113
Tool	0.074	0.079	0.061	0.068

Abbreviation: M = Mean; SD = standard deviation.

5.2.1 | Error Rates

Table 5 shows that the racial bias pattern in observed error rates descriptively replicates in all three instruction groups. We calculated a 3 (instruction) × 2 (prime) × 2 (target) × 2 (base-rate) mixed-design repeated-measures ANOVA of the error rates (for the complete ANOVA tables, see Appendix A2 and A2). Importantly, instructions significantly affected weapon identification bias. In the full ANOVA, the three-way interaction instruction × prime × target, $F(2, 508) = 5.41$, $p = 0.005$, $\eta_p^2 = 0.021$ was significant. To decompose this interaction, we conducted simple-effects analyses resulting in significant prime × target interactions for all three groups, control group, $F(1, 508) = 9.84$, $p = 0.002$, $\eta_p^2 = 0.019$; use race group, $F(1, 508) = 55.55$, $p < 0.001$, $\eta_p^2 = 0.099$; and avoid race group ($F(1, 508) = 14.08$, $p < 0.001$, $\eta_p^2 = 0.027$). Post-hoc group comparisons revealed

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Goodness-of-fit statistics and model selection indices for Study 3.

Parameter constraints	Control				Use race				Avoid race			
	$p(T_1)$	$p(T_2)$	DIC	WAIC	$p(T_1)$	$p(T_2)$	DIC	WAIC	$p(T_1)$	$p(T_2)$	DIC	WAIC
Unconstrained	0.51	0.33	5659.1	5577.7	0.43	0.06	5988.5	5868.1	0.43	0.16	5798.9	5672.5
$A_{Bmg} = A_{Wmg}$; $A_{Bmt} = A_{Wmt}$	0.58	0.34	5650.2	5583.8	0.29	< 0.001	6088.4	6045.2	0.48	0.14	5787.7	5684.3
$C_{BG} = C_{BT} = C_{WG} = C_{WT}$	0.26	0.04	5667.6	5632.6	0.36	0.002	6075.3	6073.5	0.02	0.08	5880.7	5864.1
$C_{BT} = s_{BI} \cdot C_{BG}$; $C_{WT} = s_{W1} \cdot C_{WG}$	0.40	0.33	5666.3	5602.1								
$s_{BI} = s_{W1}$	0.25	0.22	5654.7	5585.0								
$s_{BI} \leq s_{W1}$	0.25	0.23	5659.1	5589.4								
$C_{BT} = s_{B1} \cdot C_{BG}$; $C_{WG} = s_{W2} \cdot C_{WT}$					0.50	0.02	6051.9	6024.7				
$C_{BT} = s_{B1} \cdot C_{BG}$; $C_{WG} = C_{WT}$					0.50	0.02	6053.0	6023.5				
$C_{BT} = C_{BG}$; $C_{WG} = s_{W2} \cdot C_{WT}$					0.49	0.02	6052.4	6023.4				
$C_{BG} = s_{B2} \cdot C_{BT}$; $C_{WG} = s_{W2} \cdot C_{WT}$									0.38	0.09	5824.5	5737.0
$s_{B2} = s_{W2}$									0.45	0.08	5817.7	5730.3
$s_{W2} \leq s_{B2}$									0.41	0.08	5819.1	5729.8

Abbreviations: A = response tendency parameter, B = Black prime, C = successful discrimination parameter, G = gun target, mg = more guns target condition, mt = more tools target condition, s = shrinkage parameter, T = tool target, W = White prime.

that the weapon identification bias was more pronounced in the use race group compared to the control group, $F(1, 339) = 7.86$, $p = 0.005$, $\eta_p^2 = 0.023$ and the avoid race group, $F(1, 339) = 5.11$, $p = 0.024$, $\eta_p^2 = 0.015$, but did not differ between the control and avoid race group, $F(1, 338) = 0.47$, $p = 0.492$, $\eta_p^2 = 0.001$. Thus, instructions to engage in racial profiling produce a more pronounced weapon identification bias compared to instructions to avoid racial profiling or to neutral instructions.

5.2.2 | PDP Modeling

The unconstrained generalized PDP model showed good fit for all instruction groups (see Table 6). We therefore use this model as a baseline model and report parameter estimates accordingly. Again, as in Studies 1 and 2, to evaluate racial bias effects on A- and C-parameters, we compared model fit and model selection measures between the unconstrained model and models equating A- or C-parameters indicative of racial bias (Table 6).

Response bias: As illustrated in Figure 6, in the control group, mean A-parameter estimates for Black primes exceeded those of White primes for both the mg (A_{Bmg} = 0.64 [0.56–0.72]; A_{Wmg} = 0.60 [0.52–0.68]) and the mt (A_{Bmt} = 0.34 [0.27–0.42]; A_{Wmt} = 0.30 [0.24–0.37]) conditions. As in Study 1 and 2, model comparison revealed a better fit of the unconstrained model compared to the constrained model with equated A-parameters across prime conditions within base-rate conditions regarding WAIC but not DIC. Thus, the effect of race on response bias in the control group is somewhat weaker than it was in Studies 1 and 2.

In the use race group, mean A-parameter estimates for Black primes resembled those for White primes in the mg (A_{Bmg} = 0.63 [0.53–0.73]; A_{Wmg} = 0.63 [0.53–0.71]) condition and were only slightly higher in the mt (A_{Bmt} = 0.35 [0.27–0.44]; A_{Wmt} = 0.34 [0.26–0.43]) condition. Nevertheless, model comparison showed slightly better fit of the unconstrained model compared to the constrained model with equal A-parameters across prime conditions within base-rate conditions, suggesting a small effect of race on response bias in the use race group.

In the avoid race group, mean A-parameter estimates for Black primes were descriptively higher compared to White primes for both the mg (A_{Bmg} = 0.66 [0.54–0.77]; A_{Wmg} = 0.64 [0.53–0.74]) and the mt (A_{Bmt} = 0.39 [0.28–0.52]; A_{Wmt} = 0.35 [0.26–0.47]) condition. Resembling the pattern in the control condition, model comparison showed a better model fit for the unconstrained model compared to the constrained model with equal A-parameters regarding WAIC but not DIC. Thus, the evidence for an influence of race on response bias in the avoid race group is somewhat mixed, comparable to the control group condition.

Discrimination bias: As illustrated in Figure 7, in the control group, mean C-parameter estimates for gun targets were descriptively slightly higher compared to tool targets in the Black prime condition (C_{BG} = 0.84 [0.81–0.87]; C_{BT} = 0.81 [0.76–0.85]) but equal in the White prime condition (C_{WG} = 0.83 [0.80–0.86]; C_{WT} = 0.83 [0.79–0.87]). Again, to test for a prime-target interaction, we compared the model fit of a baseline model assuming $C_{BT} = s_{BI} \cdot C_{BG}$ and $C_{WT} = s_{W1} \cdot C_{WG}$ to a model with the same parameter structure and the additional equality constraint $s_{BI} = s_{W1}$. The comparison revealed that the model with the equality constraint ($s_{BI} = s_{W1}$) fits the data better than both the model without this constraint (s_{BI}; s_{W1}) and the ordinal prime-target interaction model ($s_{BI} \leq s_{W1}$). As in Study 1 and the control group of Study 2, there is again no support for the discrimination bias hypothesis according to which racial bias is mediated by object discrimination. Also, the data again suggest a generally higher probability of successful discrimination for gun compared to tool targets, irrespective of prime race. In line with this, a model constraining all C-parameters to be equal (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$) fit the data worse than the model with the equality constraint ($s_{BI} = s_{W1}$).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

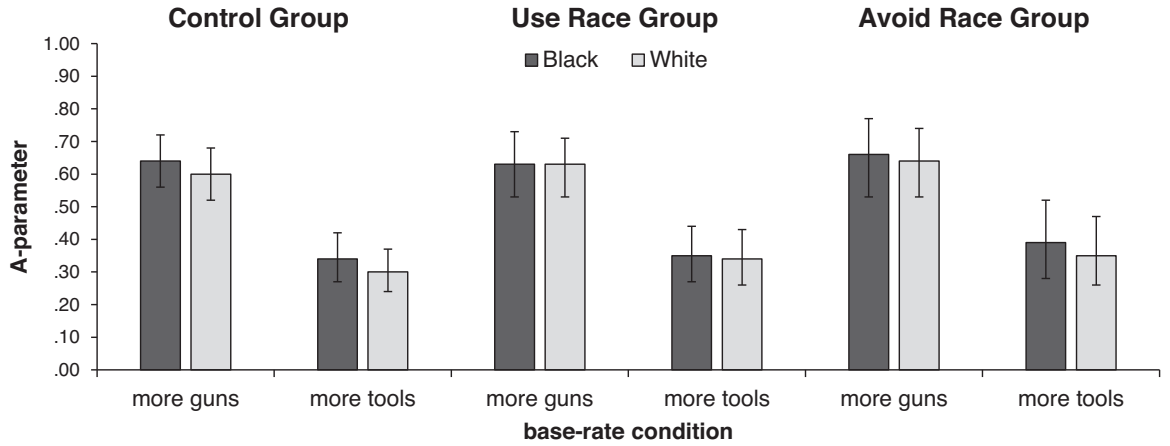

FIGURE 6 | *A*-parameter estimates of Study 3 by prime race, base-rate condition and instruction condition. *Note:* Mean parameter estimates with 95% BCIs as error bars.

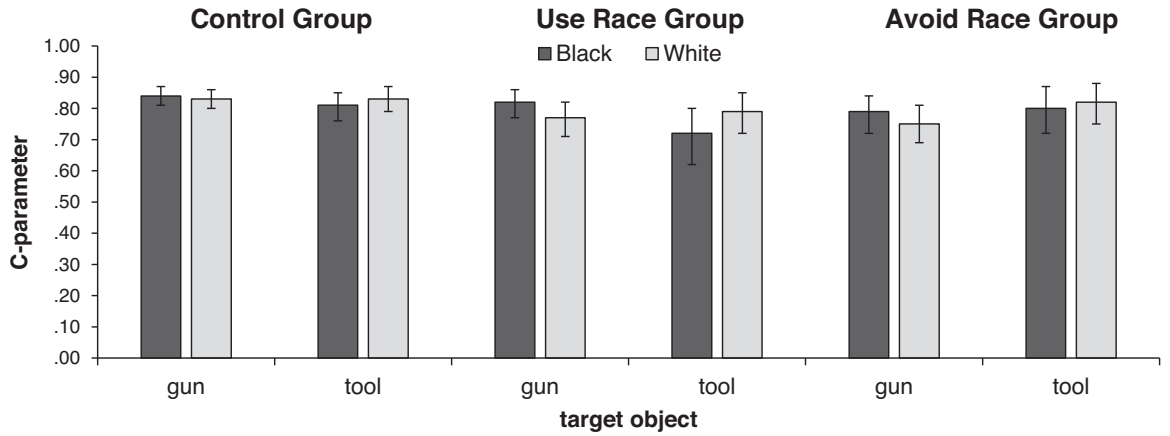

FIGURE 7 | *C*-parameter estimates of Study 3 by prime race, target object and instruction condition. *Note:* Mean parameter estimates with 95% BCIs as error bars.

In the use race group, mean *C*-parameter estimates for gun targets were descriptively higher compared to tool targets in the Black prime condition (C_{BG} = 0.82 [0.77–0.86]; C_{BT} = 0.72 [0.62–0.80]) and reversed in the White prime condition (C_{WG} = 0.77 [0.71–0.82]; C_{WT} = 0.79 [0.72–0.85]). Hence, the estimates display the expected disordinal interaction pattern (i.e., $C_{BG} > C_{BT}$ and $C_{WG} < C_{WT}$), resembling the use race group in Study 2. To test this interaction, we first compared the disordinal interaction model $C_{BT} = s_{B1} \cdot C_{BG}$ and $C_{WG} = s_{W2} \cdot C_{WG}$ to a model equating all *C*-parameters (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$). The disordinal interaction model fitted the data better than the model equating all *C*-parameters. Resembling our strategy in Study 2, we additionally estimated a model assuming a discrimination advantage restricted to guns in the Black prime condition (i.e., $C_{BT} = s_{B1} \cdot C_{BG}$ and $C_{WT} = C_{WG}$) and a model assuming a discrimination advantage restricted to tools in the White prime condition (i.e., $C_{WG} = s_{W2} \cdot C_{WG}$ and $C_{BT} = C_{BG}$). Neither the former nor the latter model provides evidence of better model fit compared to the disordinal interaction model. Replicating Study 2, this again suggests that the use race instruction biases target discrimination in both prime conditions. Overall, these results support the discrimination bias hypothesis for the use race instruction group.

In the avoid race group, mean *C*-parameter estimates for gun targets were descriptively smaller compared to tool targets in both the Black prime condition (C_{BG} = 0.79 [0.72–0.84]; C_{BT} = 0.80 [0.72–0.87]) and the White prime condition (C_{WG} = 0.75 [0.69–0.81]; C_{WT} = 0.82 [0.75–0.88]). To test for a prime-target interaction effect on *C*-parameters, we compared model fit of a baseline model assuming $C_{BG} = s_{B2} \cdot C_{BT}$ and $C_{WG} = s_{W2} \cdot C_{WT}$ to a model with the additional equality constraint $s_{B2} = s_{W2}$. Note that the former model represents an ordinal interaction opposite in sign to the effects observed in Study 1 and the control group of Studies 2 and 3, thus representing a discrimination advantage of tools over guns for the White compared to the Black prime condition. This revealed that the model with the equality constraint ($s_{B2} = s_{W2}$) fits the data better than the model without this constraint (s_{B2}; s_{W2}) and about as good as the model that implements the prime-target interaction ($s_{B2} \leq s_{W2}$). Hence, in the avoid race instruction group, there is no support for the discrimination bias hypothesis. The data rather suggest a generally higher probability of successful discrimination for tool compared to gun targets, irrespective of prime race. In line with this, a model constraining all *C*-parameters to be equal (i.e., $C_{BG} = C_{BT} = C_{WG} = C_{WT}$) does not fit the data well.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.3 | Discussion

A weapon identification bias modulated by race was observed across both base-rate conditions for all three instruction groups, but with a more pronounced weapon identification bias when participants engaged in racial profiling. The modeling results for the control and use race group replicated corresponding results of Studies 1 and 2.

Across all instruction groups, *A*-parameters reflected response bias effects of race. However, model comparisons were less convincing for the response bias hypothesis than in the first two studies. This is perhaps due to the more liberal response deadlines in Study 3. Longer response deadlines are typically associated with less errors in the WIT overall (Payne 2001; Payne et al. 2002; Rivers 2017), an effect that might diminish response bias effects in error rates (Klauer and Voss 2008; Laukenmann et al. 2023).

As in the use race group of Study 2, explicit instructions to engage in racial profiling in the WIT fostered discrimination bias as an additional determinant of weapon identification bias. In contrast, in the avoid race group overall weapon identification bias was weaker, and no contribution of discrimination bias was observed. Thus, merely increasing the salience of 'racial profiling' via instructions does not suffice to trigger discrimination bias. Object discrimination only contributes to racially biased responding when individuals actively engage in a racial profiling strategy, as in the use race group.

Comparing the avoid race with the control group, we see that both groups differ in their overall pattern of target discrimination. In the control group, participants were better at discriminating guns relative to tools, while in the avoid race group participants were better in discriminating tools relative to guns. This likely reflects a general conservative shift of the decision criterion, facilitating detection of tools relative to guns when instructed to avoid use of racial cues.

Overall, these results again suggest that response bias reflected in the *A*-parameter is the major determinant of racial bias and always involved in the WIT, irrespective of instructions. According to our results, biased target discrimination may act as an additional determinant of WIT performance, but only if participants are explicitly instructed to engage in racial profiling. This aligns well with previous research showing that instructions to engage in prejudiced responding may affect discrimination in racial bias tasks (Calanchini et al. 2013, 2021).

6 | General Discussion

The current research examined the source of racial bias as typically observed in the WIT. Overall, results suggest a dominant role of response bias, a conclusion that is congruent with previous WIT research (Klein and Todd 2024; Payne et al. 2005; Todd et al. 2021). We also found that discrimination bias may add to racial bias in the WIT (Klauer et al. 2015). However, in our research, this happened only when participants were explicitly instructed to engage in racial profiling for target identification.

6.1 | The Role of Response Bias

The results of all three studies indicate that response bias is a major determinant of weapon identification bias in the WIT. This aligns well with previous findings for the WIT (Laukenmann et al. 2023; Payne et al. 2005; Todd et al. 2021) using alternative model-based approaches. For example, according to the diffusion model analysis of Todd et al. (2021), participants have a higher initial response tendency toward gun responses following Black faces. Laukenmann et al. (2023) used response time-extended MPT models to compare different process models based on the PDP. Their model comparison favored the Default Interventionist Model (DIM) (Klauer and Voss 2008) according to which an initial default response moderated by primes' race is activated instantly, in line with the findings of Todd et al. (2021). In the framework of the DIM, response bias, mirrored in the *A*-parameter, only drives the response if target object identification fails or response conflict between the default response and target detection cannot be resolved in time (Klauer and Voss 2008; Laukenmann et al. 2023). In line with this, several previous findings suggest that racial bias depends on participants' cognitive and motivational abilities to exert effective cognitive control. For example, people motivated to control their prejudices (Volpert-Esmond et al. 2020) or individuals with higher inhibitory capacities (Ito et al. 2015) display overall better WIT performance. Also, participants with depleted executive resources show more racial bias (Govorun and Payne 2006). Hence, our results endorse previous research on response bias as a major mechanism driving racial bias in the WIT.

6.2 | The Role of Discrimination Bias

The results of Studies 2 and 3 suggest that, next to response bias, discrimination bias is an additional source of racial bias in the WIT whenever participants are explicitly encouraged to use the primes' race as a stereotype-informed cue for responding. Hence, if participants are inclined to use racial profiling for target discrimination, this may influence the interpretation of early chunks of information in the light of contextual cues provided by racial stereotypes. Consequently, when racial profiling is emphasized as a response strategy, participants might misperceive presented objects and create distorted perceptions, such as confusing the shaft of a screwdriver with a barrel of a gun after seeing a Black face prime (Klauer et al. 2015; Klauer and Voss 2008; Payne et al. 2005). Importantly, only the instruction to use (rather than to avoid) racial profiling fostered discrimination bias in weapon identification. This suggests that actively engaging in racial profiling triggers racially biased target discrimination selectively. In contrast, merely increasing the salience of racial profiling as a possible response strategy (i.e., by mentioning it in the instructions, as done in the avoid race instruction) does not suffice.

This notwithstanding, instructing participants to avoid racial profiling led to changes in overall target discrimination compared to the control group: While participants better discriminated gun objects in the control group, they better discriminated tool targets in the avoid race group. This might reflect a shift in attention focus or a conservative shift of the decision criterion, favoring detection of tools relative to guns under avoid race instructions, irrespective

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of primes. Overall, target discrimination was worse in both the use and avoid race groups compared to the control group. Hence, emphasizing racial stereotypes in WIT instructions has a general negative effect on overall performance, irrespective of whether racial profiling is encouraged or discouraged.

6.3 | The Possible Role of Conflict Resolution

Besides stereotype-driven perceptual processing, conflict resolution failure may provide an alternative determinant of discrimination bias. Specifically, resolving conflicts between different streams of task-relevant information may account for variations in the *C*-parameter across prime-target conditions (Ito et al. 2015; Klauer and Voss 2008; Laukenmann et al. 2023). According to the conflict resolution account, participants are more successful in combining the extracted perceptual information from the target with the default response elicited by the prime race if both pieces of information are congruent rather than incongruent. For example, a gun target following a Black face prime produces less conflict between the default gun response and the perceptual information compared to a tool target following a Black face prime. Although Study 1, the control groups in Studies 2 and 3 and the avoid race group revealed no prime-target interaction effect for the *C*-parameter, the use race groups in Studies 2 and 3 did. The conflict resolution account would explain this by an increased cognitive conflict when participants engage in racial profiling: Racial profiling counteracts the inhibition of incongruent default responses, resulting in worse target detection performance in incongruent trials and better performance in congruent trials.

Unfortunately, our generalized PDP model does not provide for a straightforward way to disentangle whether stereotype-driven perceptual processing, conflict resolution failure, or both underly discrimination bias as observed in the use race groups of Studies 2 and 3. Hence, this remains an open question for future research.

6.4 | Specification of the Process Dissociation Model

Our research is also informative with respect to the question whether the PDP in the conventionally used form illustrated in Figure 1—equating the *C*-parameters across target conditions—is a reasonable measurement model for WIT performance (Payne 2001). Employing a generalized PDP model for two base-rate conditions, we observed no interaction effect on the *C*-parameter in Study 1, in the control groups of Studies 2 and 3 and in the avoid race group of Study 3. However, a main effect of the target object was observed for these groups, with credibly larger *C*-estimates for gun than for tool targets for the control groups, and vice versa for the avoid race group. This aligns well with Klauer et al.'s (2015) observation that equating the *C*-parameter across primes (resulting in one *C*-parameter for guns and another for tools) may also be a legitimate PDP model specification besides the conventional PDP model. Thus, when only the data of the standard four WIT conditions are available (Black vs. White primes crossed with guns vs. tools), the PDP model best supported by the current evidence would include two separate *C*-parameters for guns and tools next to two separate *A*-parameters for Black and White face primes. To be on the safe side, however, it may be advantageous to use generalized PDP models in future research that allow the *C*-parameter to vary freely between all prime race and target type conditions (i.e., model M1 of Klauer et al. 2015, which is also the model we used). As our Studies 2 and 3 showed, this is particularly important when experimental manipulations are used that foster discrimination bias in addition to response bias.

6.5 | Limitations

Three possible limitations of the present research should be discussed briefly. First, our generalized PDP model relies on the assumption that target base-rates affect the response tendency *A* selectively, that is, successful discrimination (the *C*-parameter in the PDP) is not affected by the base-rates of targets. While this assumption might seem debatable, it is relatively common in cognitive research and has previously been employed in research using the PDP, other MPT models or signal detection theory (e.g., Buchner et al. 1995; Conrey et al. 2005; Klauer et al. 2003). We thus see no reason why this assumption should be more of a problem for our base-rate manipulation than for the pay-off manipulation previously used by Klauer et al. (2015) to manipulate *A*-parameters selectively. Moreover, possible effects of base-rates beyond those on *A* would only be a serious problem if they affect the critical prime race × target object interaction that reflects weapon identification bias. Notably, both our accuracy and response time data did not provide evidence that base-rates modulate weapon identification bias reliably. Across studies and conditions, the three-way interaction between base-rate, target object and prime race was not significant in almost all 12 tests, neither for error rates (except for error rates in the control group of Study 2) nor for correct response times (see Appendices A2 and A4). These findings suggest that the base-rate manipulation did not systematically influence the critical prime race × target object interaction and is therefore unlikely a source of bias in our analyses.

Second, the base-rate manipulation produced generally higher error rates for low base-rate targets compared to high base-rate targets. One might suspect that this asymmetry counteracts weapon identification bias. Crucially, despite this difference in accuracy for low and high base-rate targets, the error rates demonstrated the robust prime-target interaction indicative of racial bias in all studies. In line with this, the base-rate manipulation is mirrored in the size of the *A*-parameters of the PDP as intended, that is, with large (small) *A*-parameters in the mg (mt) base-rate condition. This shows that the base-rate asymmetry did not eliminate participants' racial bias in weapon identification. Racial bias in the WIT emerged in all studies and for all base-rate conditions.

Third, the short response time window in Studies 1 and 2 led to relatively high error rates, especially for low base-rate targets. This may appear problematic as it places boundaries on object discrimination success reflected in the *C*-parameter. However, this is not an issue for PDP applications if the sum of *gun* and *tool* error rates does not exceed 1. For the standard PDP-model (Figure 1), this follows directly from the estimation equation of the *C*-parameter: $C = 1 - P(\text{error} \mid \text{gun}) - P(\text{error} \mid \text{tool})$. Indeed, summed error rates exceeding 1 would lead to inadmissible, negative *C*-parameters. Importantly, this never occurred in our

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

studies (all sums ≤ 0.705), mirrored in adequate model fit of the generalized PDP model for all studies. Notably, our resulting *C*-parameters, although small in Studies 1 and 2, were reliably larger than zero, indicating that successful discrimination occurred. Finally, to address possible concerns about high error rates in Studies 1 and 2, we conducted Study 3 with a longer response time window. This resulted in substantially lower error rates as expected, while core results of Studies 1 and 2 were successfully replicated. Thus, our conclusions are not affected by the general level of errors rates nor by their asymmetry across base rate conditions.

7 | Conclusion

In three independent pre-registered studies based on $N = 1.015$ participants, we found support for the notion that response bias is the main driver of racial bias in the WIT. This means that stereotypic information, elicited by the primes' race, generates an additional, interfering stream of information besides perceptual object discrimination. It fosters the typical racially biased response pattern in weapon identification whenever object discrimination fails within the time window provided for responding.

This notwithstanding, if participants use racial profiling as a strategy for object identification, discrimination bias additionally fosters racially biased responding. Discrimination bias results from stereotype-driven context cues elicited by racial primes when participants aim to resolve the perceptual ambiguity in target object identification or from conflict resolution failure when target objects and default responses are incongruent. In general, therefore, both mechanisms can lead to racially biased responding in weapon identification, although discrimination bias becomes important only when people engage in racial profiling.

From an applied perspective, our results suggest at least two strategies to weaken racial biases in spontaneous behavior: First, reducing response bias distortions by enhancing the capacity to inhibit stereotypic information elicited by racial primes, and second, facilitating unbiased object discrimination by preventing racial profiling strategies.

Acknowledgements

The authors are very grateful to C.E. Phills and A.M. Rivers for sharing the stimulus materials we used for our studies. The authors acknowledge the use of ChatGPT GPT-5.5 (OpenAI) and Claude Opus 4.8 and Fable 5 (Anthropic) accessed in July and August 2026 following manuscript preparation by the authors to support detection of typographical errors and to assist with final text editing, analytic code creation, figure and table creation and literature review for the revised version of our manuscript. All output was critically evaluated and edited by the authors. The responsibility for the content rests solely with the authors.

Open access funding enabled and organized by Projekt DEAL.

Funding

Manuscript preparation and data collection were supported by the Research Training Group 'Statistical Modeling in Psychology' (SMiP), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (GRK 2277—Projektnummer 310365261), and by a Doctoral Research Fellowship from the Stiftung der Deutschen Wirtschaft gGmbH (sdw, Foundation of German Business), funded by the German Federal Ministry of Education and Research (awarded to R.L.).

Ethics Statement

Ethic approval was not required for this study in accordance with local legislation and institutional requirements.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in the Open Science Framework at: https://osf.io/4yxv9. Online Supporting Information, including the pre-registration of all studies reported, are also available at: https://osf.io/4yxv9.

Endnotes

[1] This analysis strategy deviates slightly from the pre-registered analysis plan for Studies 1 and 2. The pre-registered plan aimed to compare the goodness-of-fit of different restriction patterns that represent the two hypotheses of interest (discrimination bias vs. response bias, respectively; see Klauer et al. 2015). In contrast, the analysis procedure applied here assessed parameter estimates for the generalized PDP model directly to check for the third possibility that racial bias is mediated by both response bias and discrimination bias simultaneously.

[2] To check the robustness of the results against different modeling approaches, we also estimated each model of Study 1 to 3 using maximum likelihood estimation based on pooled data (i.e., aggregated across participants). Results are listed in the Supporting Information. Importantly, nested model comparisons showed the same order of model preferences as the Bayesian approach, thus confirming robustness.

[3] In line with Cohen (1988), boldface notation indicates population effect sizes (e.g., $\boldsymbol{\eta_p}^2$) whereas standard notation indicates estimated sample effect sizes (e.g., $\eta_p{}^2$) in what follows.

[4] Note that 95% Bayesian credible intervals (BCIs) are indicated in brackets.

[5] The final sample size per group was thus slightly less than the intended 170 participants per group. This was due to a coding error in identifying Tukey's criterion for missing responses during data collection, leading to a premature stop of data collection. Tukey's criterion was slightly lower for the final data set, leading to the exclusion of more participants.

References

Bishara, A. J., and B. K. Payne. 2009. "Multinomial Process Tree Models of Control and Automaticity in Weapon Misidentification." *Journal of Experimental Social Psychology* 45, no. 3: 524–534. https://doi.org/10.1016/j.jesp.2008.11.002.

Buchner, A., E. Erdfelder, and B. Vaterrodt-Plünnecke. 1995. "Toward Unbiased Measurement of Conscious and Unconscious Memory Processes Within the Process Dissociation Framework." *Journal of Experimental Psychology: General* 124, no. 2: 137–160. https://doi.org/10.1037/0096-3445.124.2.137.

Burnham, K. P., and D. R. Anderson. 2004. "Multimodel Inference: Understanding AIC and BIC in Model Selection." *Sociological Methods & Research* 33, no. 2: 261–304. https://doi.org/10.1177/0049124104268644.

Calanchini, J., K. Gonsalkorale, J. W. Sherman, and K. C. Klauer. 2013. "Counter-Prejudicial Training Reduces Activation of Biased Associa-

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tions and Enhances Response Monitoring." *European Journal of Social Psychology* 43, no. 5: 321–325. https://doi.org/10.1002/ejsp.1941.

Calanchini, J., C. K. Lai, and K. C. Klauer. 2021. "Reducing Implicit Racial Preferences: III. A Process-Level Examination of Changes in Implicit Preferences." *Journal of Personality and Social Psychology* 121, no. 4: 796–818. https://doi.org/10.1037/pspi0000339.

Clark-Carter, D. 2004. *Quantitative Psychological Research: A Student's Handbook*. Psychology Press.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum.

Conrey, F. R., J. W. Sherman, B. Gawronski, K. Hugenberg, and C. J. Groom. 2005. "Separating Multiple Processes in Implicit Social Cognition: The Quad Model of Implicit Task Performance." *Journal of Personality and Social Psychology* 89, no. 4: 469–487. https://doi.org/10.1037/0022-3514.89.4.469.

De Houwer, J. 2014. "A Propositional Model of Implicit Evaluation." *Social and Personality Psychology Compass* 8, no. 7: 342–353. https://doi.org/10.1111/spc3.12111.

Desmond, M., A. V. Papachristos, and D. S. Kirk. 2016. "Police Violence and Citizen Crime Reporting in the Black Community." *American Sociological Review* 81, no. 5: 857–876. https://doi.org/10.1177/0003122416663494.

Edwards, F., H. Lee, and M. Esposito. 2019. "Risk of Being Killed by Police Use of Force in the United States by Age, Race–Ethnicity, and Sex." *Proceedings of the National Academy of Sciences* 116, no. 34: 16793–16798.

Erdfelder, E., T. S. Auer, B. E. Hilbig, A. Aßfalg, M. Moshagen, and L. Nadarevic. 2009. "Multinomial Processing Tree Models: A Review of the Literature." *Zeitschrift für Psychologie/Journal of Psychology* 217, no. 3: 108–124.

Faul, F., E. Erdfelder, A. G. Lang, and A. Buchner. 2009. "Statistical Power Analysis Using G*Power 3.1: Tests for Correlation and Regression Analyses." *Behavior Research Methods* 41: 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Gawronski, B., and G. V. Bodenhausen. 2006. "Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change." *Psychological Bulletin* 132, no. 5: 692–731. https://doi.org/10.1037/0033-2909.132.5.692.

Glaser, J. 2015. *Suspect Race: Causes and Consequences of Racial Profiling*. Oxford University Press.

Gonsalkorale, K., J. W. Sherman, and K. C. Klauer. 2009. "Aging and Prejudice: Diminished Regulation of Automatic Race Bias Among Older Adults." *Journal of Experimental Social Psychology* 45, no. 2: 410–414. https://doi.org/10.1016/j.jesp.2008.11.004.

Govorun, O., and B. K. Payne. 2006. "Ego-Depletion and Prejudice: Separating Automatic and Controlled Components." *Social Cognition* 24, no. 2: 111–136. https://doi.org/10.1521/soco.2006.24.2.111.

Heck, D. W., N. R. Arnold, and D. Arnold. 2018. "TreeBUGS: An R Package for Hierarchical Multinomial-Processing-Tree Modeling." *Behavior Research Methods* 50, no. 1: 264–284. https://doi.org/10.3758/s13428-017-0869-7.

Higgins, E. T. 1996. "Activation: Accessibility, and Salience." In *Social Psychology: Handbook of Basic Principles*, edited by E. T. Higgins and A. W. Kruglanski, 133–168. The Guilford Press.

Huntsinger, J. R., S. Sinclair, and G. L. Clore. 2009. "Affective Regulation of Implicitly Measured Stereotypes and Attitudes: Automatic and Controlled Processes." *Journal of Experimental Social Psychology* 45, no. 3: 560–566. https://doi.org/10.1016/j.jesp.2009.01.007.

Ito, T. A., N. P. Friedman, B. D. Bartholow, et al. 2015. "Toward a Comprehensive Understanding of Executive Cognitive Function in Implicit Racial Bias." *Journal of Personality and Social Psychology* 108, no. 2: 187–218. https://doi.org/10.1037/a0038557.

Jacoby, L. L. 1991. "A Process Dissociation Framework: Separating Automatic From Intentional Uses of Memory." *Journal of Memory and Language* 30, no. 5: 513–541. https://doi.org/10.1016/0749-596X(91)90025-F.

Klauer, K. C. 2010. "Hierarchical Multinomial Processing Tree Models: A Latent-Trait Approach." *Psychometrika* 75, no. 1: 70–98. https://doi.org/10.1007/s11336-009-9141-0.

Klauer, K. C., K. Dittrich, C. Scholtes, and A. Voss. 2015. "The Invariance Assumption in Process-Dissociation Models: An Evaluation Across Three Domains." *Journal of Experimental Psychology: General* 144, no. 1: 198–221. https://doi.org/10.1037/xge0000044.

Klauer, K. C., J. Mierke, and J. Musch. 2003. "The Positivity Proportion Effect: A List Context Effect in Masked Affective Priming." *Memory & Cognition* 31, no. 6: 953–967.

Klauer, K. C., and A. Voss. 2008. "Effects of Race on Responses and Response Latencies in the Weapon Identification Task: A Test of Six Models." *Personality and Social Psychology Bulletin* 34, no. 8: 1124–1140. https://doi.org/10.1177/0146167208318603.

Klein, S. A., and A. R. Todd. 2024. "Emotion Expression Salience and Racially Biased Weapon Identification: A Diffusion Modeling Approach." *Psychonomic Bulletin & Review* 31: 2767–2775. https://doi.org/10.3758/s13423-024-02526-z.

Klein, S. A. W., B. Lassetter, R. Neel, and A. R. Todd. 2026. "Integrating Body Information With Faces Directs Attention Away From Race, Altering Racially Biased Weapon Identification." *Psychonomic Bulletin & Review* 33, no. 1: 30. https://doi.org/10.3758/s13423-025-02802-6.

Knapp, B. R., and W. H. Batchelder. 2004. "Representing Parametric Order Constraints in Multi-Trial Applications of Multinomial Processing Tree Models." *Journal of Mathematical Psychology* 48, no. 4: 215–229. https://doi.org/10.1016/j.jmp.2004.03.002.

Kuhlmann, B. G., E. Erdfelder, and M. Moshagen. 2019. "Testing Interactions in Multinomial Processing Tree Models." *Frontiers in Psychology* 10: 2364. https://doi.org/10.3389/fpsyg.2019.02364.

Lai, C. K., and J. A. Lisnek. 2023. "The Impact of Implicit-Bias-Oriented Diversity Training on Police Officers' Beliefs, Motivations, and Actions." *Psychological Science* 34, no. 4: 424–434.

Lai, C. K., M. Marini, S. A. Lehr, et al. 2014. "Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions." *Journal of Experimental Psychology: General* 143, no. 4: 1765–1785. https://doi.org/10.1037/a0036260.

Laukenmann, R., E. Erdfelder, D. W. Heck, and M. Moshagen. 2023. "Cognitive Processes Underlying the Weapon Identification Task: A Comparison of Models Accounting for Both Response Frequencies and Response Times." *Social Cognition* 41, no. 2: 137–164. https://doi.org/10.1521/soco.2023.41.2.137.

Paluck, E. L., R. Porat, C. S. Clark, and D. P. Green. 2021. "Prejudice Reduction: Progress and Challenges." *Annual Review of Psychology* 72, no. 1: 533–560. https://doi.org/10.1146/annurev-psych-071620-030619.

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, no. 2: 181–192. https://doi.org/10.1037/0022-3514.81.2.181.

Payne, B. K. 2005. "Conceptualizing Control in Social Cognition: How Executive Functioning Modulates the Expression of Automatic Stereotyping." *Journal of Personality and Social Psychology* 89, no. 4: 488–503. https://doi.org/10.1037/0022-3514.89.4.488.

Payne, B. K., and J. Correll. 2020. "Race, Weapons, and the Perception of Threat." *Advances in Experimental Social Psychology* 62: 1–50. https://doi.org/10.1016/bs.aesp.2020.04.001.

Payne, B. K., A. J. Lambert, and L. L. Jacoby. 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, no. 4: 384–396. https://doi.org/10.1016/S0022-1031(02)00006-9.

Payne, B. K., Y. Shimizu, and L. L. Jacoby. 2005. "Mental Control and Visual Illusions: Toward Explaining Race-Biased Weapon Misidentifica-

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tions." *Journal of Experimental Social Psychology* 41, no. 1: 36–47. https://doi.org/10.1016/j.jesp.2004.05.001.

Phills, C. E., K. Kawakami, E. Tabi, D. Nadolny, and M. Inzlicht. 2011. "Mind the Gap: Increasing Associations Between the Self and Blacks With Approach Behaviors." *Journal of Personality and Social Psychology* 100, no. 2: 197–210. https://doi.org/10.1037/a0022159.

Pierson, E., C. Simoiu, J. Overgoor, et al. 2020. "A Large-Scale Analysis of Racial Disparities in Police Stops Across the United States." *Nature Human Behaviour* 4, no. 7: 736–745. https://doi.org/10.1038/s41562-020-0858-1.

Plant, E. A., and B. M. Peruche. 2005. "The Consequences of Race for Police Officers' Responses to Criminal Suspects." *Psychological Science* 16, no. 3: 180–183. https://doi.org/10.1111/j.0956-7976.2005.00800.x.

Rivers, A. M. 2017. "The Weapons Identification Task: Recommendations for Adequately Powered Research." *PLoS ONE* 12, no. 6: e0177857. https://doi.org/10.1371/journal.pone.0177857.

Röhner, J., and C. K. Lai. 2021. "A Diffusion Model Approach for Understanding the Impact of 17 Interventions on the Race Implicit Association Test." *Personality and Social Psychology Bulletin* 47, no. 9: 1374–1389.

Schmidt, O., E. Erdfelder, and D. W. Heck. 2025. "How to Develop, Test, and Extend Multinomial Processing Tree Models: A Tutorial." *Psychological Methods* 30, no. 4: 720–743. https://doi.org/10.1037/met0000561.

Sierra-Arévalo, M. 2021. "American Policing and the Danger Imperative." *Law & Society Review* 55, no. 1: 70–103. https://doi.org/10.1111/lasr.12526.

Sierra-Arévalo, M., and J. Nix. 2020. "Gun Victimization in the Line of Duty: Fatal and Nonfatal Firearm Assaults on Police Officers in the United States, 2014–2019." *Criminology & Public Policy* 19, no. 3: 1041–1066. https://doi.org/10.1111/1745-9133.12507.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. V. D. Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64, no. 4: 583–639. https://doi.org/10.1111/1467-9868.00353.

Todd, A. R., D. J. Johnson, B. Lassetter, R. Neel, A. J. Simpson, and J. Cesario. 2021. "Category Salience and Racial Bias in Weapon Identification: A Diffusion Modeling Approach." *Journal of Personality and Social Psychology* 120, no. 3: 672–693. https://doi.org/10.1037/pspi0000279.

Vehtari, A., A. Gelman, and J. Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." *Statistics and Computing* 27, no. 5: 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.

Volpert-Esmond, H. I., L. D. Scherer, and B. D. Bartholow. 2020. "Dissociating Automatic Associations: Comparing Two Implicit Measurements of Race Bias." *European Journal of Social Psychology* 50, no. 4: 876–888. https://doi.org/10.1002/ejsp.2655.

Watanabe, S. 2010. "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory." *Journal of Machine Learning Research* 11: 3571–3594.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information Table S1. Overview of parameter estimates of Study 1, 2, and 3 **Supporting Information Table S2**. Model fit of pooled data MPT-model Study 1 **Supporting Information Table S3**. Model fit of pooled data MPT-model Study 2 **Supporting Information Table S4.1 + S4.2**. Model fit of pooled data MPT-model Study 3 **Supporting Information Table S5**. Model fit of preregistered and exploratory analysis Study 1 and 2

Appendix

Appendix 1: Racial Profiling Instructions for Study 2 and Study 3

Appendix 1a: Use Race Instruction of Study 2 and Study 3 (Payne et al. 2002, 388)

'The faces will be from either White (European American) or Black (African American) people. Research has shown that the race of the face sometimes impacts the ways that people classify the second object. People are sometimes faster and more accurate in responding to guns after a Black face than after a White face. You have been randomly assigned to take the "racial profiling" condition. Regardless of your personal views, we would like you to play the role of someone engaged in racial profiling. That is, try to make correct classifications, but we would like you to use the race of the faces to help you identify the gun or tool in question.'

Appendix 1b: Avoid Race Instruction of Study 3 (Payne et al. 2002, 388)

'The faces will be from either White (European American) or Black (African American) people. Research has shown that the race of the face sometimes impacts the ways that people classify the second object. People are sometimes faster and more accurate in responding to guns after a Black face than after a White face. You have been randomly assigned to take the perspective of a completely unbiased person. Regardless of your personal views, we would like you to base your responses only on whether the second object looks more like a gun or tool. Try not to let the race of the face influence your decisions.'

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix 2: ANOVAs of Error Rates of Studies 1, 2 and 3

TABLE A2a | ANOVA for error rates in Study 1.

Predictor	$F(1, 169)$	p	η^2_p
Intercept	1829.63	< 0.001	
Base rate condition	6.548	0.011	0.037
Prime race	1.082	0.300	0.006
Target object	0.027	0.870	< 0.001
Base rate condition × prime race	0.436	0.510	0.003
Base rate condition × target object	1026.386	< 0.001	0.859
Prime race × target object	10.900	0.001	0.061
Base rate condition × prime race × target object	1.341	0.249	0.008

Note: Repeated-measures ANOVA for error rates (Study 1).

TABLE A2b | ANOVA for error rates in Study 2.

Predictor	$F(1, 332)$	p	η^2_p
Intercept	2542.047	< 0.001	
Base rate condition	9.661	0.002	0.028
Prime race	3.496	0.062	0.010
Target object	10.192	0.002	0.030
Instruction condition	0.529	0.467	0.002
Base rate condition × prime race	18.790	< 0.001	0.054
Base rate condition × target object	1629.856	< 0.001	0.831
Base rate condition × instruction condition	1.2901	0.257	0.004
Prime race × target object	86.786	< 0.001	0.207
Prime race × instruction condition	1.979	0.160	0.006
Target object × instruction condition	0.082	0.775	< 0.001
Base rate condition × prime race × target object	8.782	0.003	0.026
Base rate condition × prime race × instruction condition	7.776	0.006	0.023
Base rate condition × target object × instruction condition	1.779	0.183	0.005
Prime race × target object × instruction condition	22.598	< 0.001	0.064
Base rate condition × prime race × target object × instruction condition	0.049	0.825	< 0.001

Note: Mixed-measures ANOVA for error rates (Study 2).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A2c | ANOVA for error rates in Study 2.

Predictor	$F(1, 332)$	p	η^2_p
Control group			
Base rate condition	9.115	0.003	0.027
Prime race	0.108	0.742	< 0.001
Target object	4.276	0.039	0.013
Base rate condition × prime race	1.210	0.272	0.004
Base rate condition × target object	880.203	< 0.001	0.726
Prime race × target object	10.533	0.001	0.031
Base rate condition × prime race × target object	5.134	0.024	0.015
Use race group			
Base rate condition	1.922	0.167	0.006
Prime race	5.304	0.022	0.016
Target object	5.977	0.015	0.018
Base rate condition × prime race	25.070	< 0.001	0.070
Base rate condition × target object	752.955	< 0.001	0.694
Prime race × target object	97.806	< 0.001	0.228
Base rate condition × prime race × target object	3.713	0.055	0.011

Note: Joint estimation of repeated-measures ANOVA for error rates with simple effects broken down by instruction group (Study 2).

TABLE A2d | ANOVA for error rates in Study 3.

Predictor	$F(1, 508)$ or $F(2, 508)$	p	η^2_p
Intercept	895.393	< 0.001	
Base rate condition	0.963	0.327	0.002
Prime race	0.537	0.464	0.001
Target object	5.103	0.024	0.010
Instruction condition	1.194	0.303	0.005
Base rate condition × prime race	6.615	0.010	0.013
Base rate condition × target object	548.659	< 0.001	0.519
Base rate condition × instruction condition	1.445	0.237	0.006
Prime race × target object	68.488	< 0.001	0.119
Prime race × instruction condition	3.497	0.031	0.014
Target object × instruction condition	2.561	0.078	0.010
Base rate condition × prime race × target object	0.261	0.610	< 0.001
Base rate condition × prime race × instruction condition	0.102	0.903	< 0.001
Base rate condition × target object × instruction condition	0.091	0.913	< 0.001
Prime race × target object × instruction condition	5.414	0.005	0.021
Base rate condition × prime race × target object × instruction condition	1.240	0.290	0.005

Note: Mixed-measures ANOVA for error rates (Study 3). The F-statistic for predictors including instruction conditions are $F(2, 508)$. All other are $F(1, 508)$.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A2e | ANOVA for error rates in Study 3.

Predictor	$F(1, 508)$	p	η^2_p
Control group			
Base rate condition	0.113	0.737	< 0.001
Prime race	3.203	0.074	0.006
Target object	3.656	0.056	0.007
Base rate condition × prime race	1.940	0.164	0.004
Base rate condition × target object	178.924	< 0.001	0.260
Prime race × target object	9.838	0.002	0.019
Base rate condition × prime race × target object	0.812	0.368	0.002
Use race group			
Base rate condition	0.011	0.917	< 0.001
Prime race	1.416	0.235	0.003
Target object	0.261	0.610	< 0.001
Base rate condition × prime race	3.394	0.067	0.007
Base rate condition × target object	193.084	< 0.001	0.275
Prime race × target object	55.550	< 0.001	0.099
Base rate condition × prime race × target object	0.234	0.629	< 0.001
Avoid race group			
Base rate condition	3.726	0.054	0.007
Prime race	2.916	0.088	0.006
Target object	6.280	0.013	0.012
Base rate condition × prime race	1.489	0.223	0.003
Base rate condition × target object	176.891	< 0.001	0.258
Prime race × target object	14.078	< 0.001	0.027
Base rate condition × prime race × target object	1.696	0.193	0.003

Note: Joint estimation of repeated-measures ANOVA for error rates with simple effects broken down by instruction group (Study 3).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix 3: Mean Correct Response Times by Prime, Target and Base-Rate of Studies 1, 2 and 3

TABLE A3a | Mean correct response time in Study 1.

	Prime			
	Black		White	
Target	***M***	**SD**	***M***	**SD**
More guns condition				
Gun	347	59	353	61
Tool	404	80	402	72
More tools condition				
Gun	388	63	391	64
Tool	358	64	351	59

Note: Mean correct response times (in ms) by prime, target and base-rate condition in Study 1.
Abbreviation: SD = standard deviation.

TABLE A3b | Mean correct response time in Study 2.

	Prime			
	Black		White	
Target	***M***	**SD**	***M***	**SD**
Control group				
More guns condition				
Gun	369	60	371	61
Tool	430	78	430	72
More tools condition				
Gun	417	57	418	67
Tool	383	60	379	57
Use race group				
More guns condition				
Gun	366	57	376	59
Tool	439	79	430	71
More tools condition				
Gun	409	74	414	71
Tool	380	67	368	64

Note: Mean correct response times (in ms) by prime, target and base-rate condition in Study 2.
Abbreviation: SD = standard deviation.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A3c | Mean correct response time in Study 3.

	Prime			
	Black		White	
Target	***M***	**SD**	***M***	**SD**
Control group				
More guns condition				
Gun	468	49	475	61
Tool	513	53	511	50
More tools condition				
Gun	488	47	493	50
Tool	499	48	494	48
Use race group				
More guns condition				
Gun	460	50	472	48
Tool	515	49	507	49
More tools condition				
Gun	481	52	493	54
Tool	498	53	488	53
Avoid race group				
More guns condition				
Gun	470	53	477	53
Tool	515	51	510	53
More tools condition				
Gun	489	49	494	50
Tool	497	51	492	50

Note: Mean correct response times (in ms) by prime, target and base-rate condition in Study 3.
Abbreviation: SD = standard deviation.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix 4: ANOVAs of Correct Response Times of Studies 1, 2 and 3

TABLE A4a | ANOVA for correct response times in Study 1.

Predictor	$F(1, 169)$	p	η^2_p
Intercept	207,650	< 0.001	
Base rate condition	0.444	0.506	0.003
Prime race	0.570	0.451	0.003
Target object	8.254	0.005	0.047
Base rate condition × prime race	5.038	0.026	0.029
Base rate condition × target object	522.200	< 0.001	0.755
Prime race × target object	6.891	0.009	0.039
Base rate condition × prime race × target object	0.334	0.564	0.002

Note: Repeated-measures ANOVA for log-transformed correct response times (Study 1).

TABLE A4b | ANOVA for correct response times in Study 2.

Predictor	$F(1, 330)$	p	η^2_p
Intercept	453,610	< 0.001	
Base rate condition	2.303	0.130	0.007
Prime race	0.268	0.605	< 0.001
Target object	56.620	< 0.001	0.146
Instruction condition	0.170	0.680	< 0.001
Base rate condition × prime race	5.453	0.020	0.016
Base rate condition × target object	869.500	< 0.001	0.725
Base rate condition × instruction condition	3.713	0.055	0.011
Prime race × target object	20.262	< 0.001	0.058
Prime race × instruction condition	0.694	0.406	0.002
Target object × instruction condition	0.197	0.658	< 0.001
Base rate condition × prime race × target object	0.334	0.564	0.001
Base rate condition × prime race × instruction condition	0.409	0.523	0.001
Base rate condition × target object × instruction condition	0.357	0.551	0.001
Prime race × target object × instruction condition	10.206	0.002	0.030
Base rate condition × prime race × target object × instruction condition	0.001	0.979	< 0.001

Note: Mixed-measures ANOVA for log-transformed correct response times (Study 2).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A4c | ANOVA for correct response times in Study 2.

Predictor	$F(1, 330)$	p	η^2_p
Control group			
Base rate condition	0.085	0.770	< 0.001
Prime race	0.050	0.822	< 0.001
Target object	25.533	< 0.001	0.072
Base rate condition × prime race	1.464	0.227	0.004
Base rate condition × target object	424.993	< 0.001	0.563
Prime race × target object	0.869	0.352	0.003
Base rate condition × prime race × target object	0.186	0.666	< 0.001
Use race group			
Base rate condition	5.826	0.016	0.017
Prime race	0.896	0.344	0.003
Target object	31.183	< 0.001	0.086
Base rate condition × prime race	4.347	0.038	0.013
Base rate condition × target object	444.514	< 0.001	0.574
Prime race × target object	29.089	< 0.001	0.081
Base rate condition × prime race × target object	0.149	0.699	< 0.001

Note: Joint estimation of repeated-measures ANOVA for log-transformed correct response times with simple effects broken down by instruction group (Study 2).

TABLE A4d | ANOVA for correct response times in Study 3.

Predictor	$F(1, 508)$ or $F(2, 508)$	p	η^2_p
Intercept	2,024,000	< 0.001	
Base rate condition	0.958	0.328	0.002
Prime race	9.485	0.002	0.018
Target object	500.420	< 0.001	0.496
Instruction condition	0.436	0.647	0.002
Base rate condition × prime race	3.200	0.074	0.006
Base rate condition × target object	705.410	< 0.001	0.581
Base rate condition × instruction condition	0.251	0.778	< 0.001
Prime race × target object	104.090	< 0.001	0.170
Prime race × instruction condition	1.813	0.164	0.007
Target object × instruction condition	1.615	0.200	0.006
Base rate condition × prime race × target object	0.949	0.330	0.002
Base rate condition × prime race × instruction condition	0.695	0.500	0.003
Base rate condition × target object × instruction condition	1.402	0.247	0.005
Prime race × target object × instruction condition	9.262	< 0.001	0.035
Base rate condition × prime race × target object × instruction condition	0.804	0.448	0.003

Note: Mixed-measures ANOVA for log-transformed correct response times (Study 3). The F-statistic for predictors including instruction conditions are $F(2, 508)$. All other are $F(1, 508)$.

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A4e | ANOVA for correct response times in Study 3.

Predictor	$F(1, 508)$	p	η^2_p
Control group			
Base rate condition	1.066	0.302	0.002
Prime race	1.984	0.160	0.004
Target object	161.802	< 0.001	0.242
Base rate condition × prime race	3.861	0.050	0.008
Base rate condition × target object	201.959	< 0.001	0.284
Prime race × target object	13.809	< 0.001	0.026
Base rate condition × prime race × target object	0.216	0.643	< 0.001
Use race group			
Base rate condition	0.392	0.532	< 0.001
Prime race	10.731	0.001	0.021
Target object	204.339	< 0.001	0.287
Base rate condition × prime race	0.133	0.716	< 0.001
Base rate condition × target object	275.978	< 0.001	0.352
Prime race × target object	88.054	< 0.001	0.148
Base rate condition × prime race × target object	0.079	0.778	< 0.001
Avoid race group			
Base rate condition	0.001	0.971	< 0.001
Prime race	0.428	0.513	< 0.001
Target object	137.719	< 0.001	0.213
Base rate condition × prime race	0.588	0.444	0.001
Base rate condition × target object	230.506	< 0.001	0.312
Prime race × target object	20.991	< 0.001	0.040
Base rate condition × prime race × target object	2.257	0.134	0.004

Note: Joint estimation of repeated-measures ANOVA for log-transformed correct response times with simple effects broken down by instruction group (Study 3).

10990992, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejsp.70119 by Universitat Mannheim, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License