Emotion Expression Salience and Racially Biased Weapon Identification: A Diffusion Modeling Approach

Samuel A. W. Klein[1*] and Andrew R. Todd[1*]

[1] Department of Psychology, University of California, Davis

This is a preprint of a manuscript that is currently in review

[*]Corresponding authors:

E-mail: sawklein@ucdavis.edu or atodd@ucdavis.edu

This research was facilitated by NSF grant BCS-1764097 (awarded to ART).

Abstract

Racial stereotypes are commonly activated by informational cues that are detectable in people's faces, with adverse consequences when applied to marginalized group members. Here, we used a sequential priming task to examine whether and how the relative salience of emotion versus race information in male face primes varying in apparent race (Black, White) and posed emotion expression (anger, happiness) shapes racial bias in weapon identification (gun vs. tool) decisions. In two experiments (N_{total} = 546) using two different manipulations of facial information salience, racial bias in weapon identification was weaker when the salience of emotion expression versus race was heightened. Process analyses using diffusion modeling tested competing accounts of the cognitive mechanism by which the salience of facial information moderates this behavioral effect. Consistent support emerged for an initial bias account, whereby 1) the decision process began closer to the "gun" response upon seeing faces of Black versus White men, and 2) this racially biased shift in the starting position was weaker when emotion versus race information was salient. We discuss these results vis-à-vis prior empirical and theoretical work on how facial information salience moderates racial bias in decision-making.

Emotion Expression Salience and Racially Biased Weapon Identification: A Diffusion Modeling Approach

Racial bias continues to pervade modern society, with adverse consequences for marginalized racial groups. Abundant experimental evidence indicates that perceivers more strongly associate Black versus White people with weapons (Payne & Correll, 2020). In the Weapon Identification Task (WIT), for example, participants are usually better (i.e., faster and more accurate) at identifying guns and worse at identifying harmless objects (e.g., tools, toys) after seeing Black versus White face primes (e.g., Amodio et al., 2004; Payne, 2001; Todd et al., 2016). This typical pattern of racial bias in the WIT is robust (see Rivers, 2017, for a meta-analysis); however, its magnitude may vary based on the salience of (i.e., the attention garnered by; Higgins, 1996) information that is detectable in the face primes. Indeed, racially biased weapon identification is weaker and sometimes eliminated when age versus race information is made more salient (Jones & Fazio, 2010; Todd et al., 2021). Granted, age is only one of many potential information sources to which perceivers might attend. In two experiments, we investigated whether attending to another source of facial information—emotion expression—likewise weakens weapon-related racial bias, relative to attending to race.

Unlike facial cues pertaining to comparatively more static social categories (e.g., age, gender, race), facial emotion expressions are dynamic and may signal a person's current affect and intentions, making them informative and motivationally relevant for perceivers (Niedenthal & Brauer, 2012; Todorov et al., 2008). The mere availability of emotion expressions (e.g., scowls vs. smiles) can affect racially biased weapon identification (Kubota & Ito, 2014), among other decisions (e.g., Raissi & Steele, 2021; Richeson & Trawalter, 2008). What remains unclear, however, is whether the *relative salience* of emotion versus race information moderates racially biased weapon identification decisions.

In addition to investigating *whether* facial information salience alters racial bias in the WIT, we use a hierarchical version of the diffusion decision model (DDM; Ratcliff et al., 2016) to examine *how* information salience shapes the cognitive processes underlying this bias. The DDM is a sequential sampling model designed to disentangle processes underlying behavior in binary decision tasks like the WIT by concurrently modeling both decisions and decision speed. It decomposes decisions into four parameters (see Table 1). We briefly describe two parameters that might explain how information salience moderates racially biased weapon identification.

Table 1

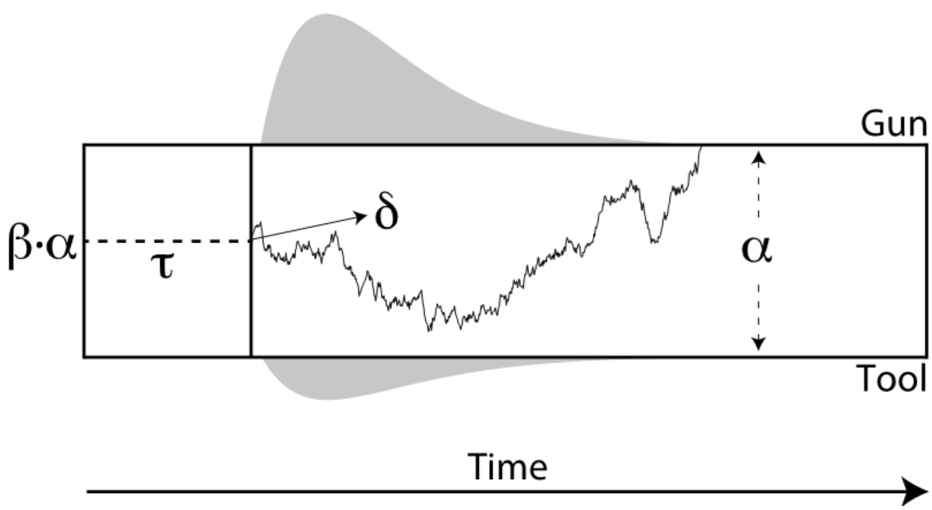
Parameters of the Diffusion Decision Model in the Weapon Identification Task

Parameter	Interpretation
Relative start point (β)	Initial bias to select *gun* or *tool* at the start of the evidence accumulation process, with 0 < β < 1. Values >.50 indicate a bias to select *gun*; values <.50 indicate a bias to select *tool*.
Threshold separation (α)	Amount of evidence required to decide, with 0 < α. Hitting a threshold triggers a decision to select *gun* or *tool*.
Drift rate (δ)	Average quality of information extracted from a stimulus at each unit of time, with -∞ < δ < ∞. Higher absolute values indicate stronger evidence. Positive values indicate evidence to select *gun*; negative values indicate evidence to select *tool*.
Non-decision time (τ)	Length of all response components (encoding time, motor response time, and other unknown contaminants) unrelated to decision making, with 0 < τ. Measured in milliseconds.

An assumption of the DDM is that decisions are made by accumulating evidence over time until reaching a decision threshold. The DDM models both the strength of evidence extracted (i.e., drift rate) and the initial position from which evidence accumulation begins (i.e., relative start point; see Figure 1). An *evidence accumulation* account of facial information salience moderating racially biased weapon identification posits that 1) seeing a Black versus White face prime strengthens the evidence accumulated for identifying guns (i.e., race-stereotypic objects), and that 2) this racial bias in evidence accumulation is weaker when emotion versus race information is salient. Alternatively, an *initial bias* account posits that 1) seeing a Black versus White face prime shifts the starting position of the decision process closer to the “gun” response, and that 2) this racially biased shift in the starting position is weaker when emotion versus race is salient.[1]

Figure 1

Illustration of the Diffusion Decision Process



Notes. People start the decision process with a bias to select gun or tool, as indicated by the relative start point, β. They then accumulate evidence (as illustrated by the jagged line) for each decision option, with average strength δ. The distance between the thresholds, α, indicates the amount of evidence needed to decide. Finally, the length of non-decision processes is indicated by τ. The hypothetical distributions (in gray) above and below the decision space indicate that the model predicts the distribution of response times for each decision option.

[1] We did not derive clear predictions about threshold separation and non-decision time, but we report results pertaining to these model parameters for completeness.

Of these two accounts, the initial bias account is more strongly supported in past work applying the DDM to the WIT.[2] Specifically, Todd et al. (2021) failed to explain racially biased weapon identification, or its moderation by the salience of age versus race information in the face primes, via an evidence accumulation account. Rather, their findings supported an initial bias account, whereby racially biased starting positions were moderated by age versus race salience. If emotion versus race salience similarly moderates racially biased weapon identification and the process underlying such decisions, then the initial bias account should emerge as the superior explanation in our experiments.

Although our focal interest was in process analyses of the DDM parameters, we also report behavioral analyses of the error rates and correct response times (RTs). For both experiments, we describe our sample size rationale, and all data exclusions, manipulations, and measures (data and code are available at https://osf.io/xdea7/).

Experiment 1

Method

Participants

Prior work using a similar design (Todd et al., 2021, Experiment 2) revealed a small-to-medium sized salience effect on racial bias in the WIT (Salience × Race Prime × Target Object interaction: $\eta_p^2 = .028$). Thus, we set a target sample size ($N = 280$) affording ≥80% power to detect $\eta_p^2 = .028$ (Faul et al., 2007). In total, 311 undergraduates consented to participate for course credit. We decided a priori to exclude data from participants who performed at or below chance (errors on ≥50% of trials) on any trial type in the WIT ($n = 21$). Retaining the excluded

[2] Notably, an evidence accumulation account better explains racial bias in the first-person shooter task (FPST, Correll et al., 2015; Johnson et al., 2018; Pleskac et al., 2018). For a discussion of procedural differences between the FPST and the WIT, see Todd et al. (2021).

data did not meaningfully alter any of the conclusions in either experiment. The final sample comprised 290 participants (81% women, 15.4% men, 1.8% non-binary; 15% White, 2.1% Black, 57.3% Asian, 17.1% Latinx, 5.2% multiracial; M_{age} = 19.3, SD = 1.3).

Procedure

In both experiments, participants arrived at the lab in small groups and were led by an experimenter to an individual computer workstation to complete the experimental tasks. Participants completed a sequential priming task, the WIT (Payne, 2001), wherein two images appeared in quick succession. Instructions urged participants to ignore the first image (face prime) and to classify the second image (target object) quickly and accurately via key press. The face primes were facial images of 48 men varying in apparent race (24 Black, 24 White) and posed emotion expression (24 angry/scowling, 24 happy/smiling) from the Chicago Face Database (Ma et al., 2015).[3] The target objects were 6 gun and 6 tool images from Payne (2001). Each trial comprised the following sequence: fixation cross (500 ms), face prime (200 ms), target object (200 ms), and pattern mask (until participants responded). If participants failed to respond within 500 ms, a message (“Please respond faster!”) appeared (1 s).

We structured the WIT so that apparent race or emotion expression was more distinctive throughout the task (Macrae & Cloutier, 2009; Rees et al., 2022; Todd et al., 2021). Participants were randomly assigned to complete one of two WIT variants, each comprising two blocks of 144 experimental trials (288 total trials that were preceded by 12 practice trials). In the *race-salient* condition, the face primes were scowling Black and White men in one block of trials and smiling Black and White men in the other block. In the *expression-salient* condition, the face

[3] The emotion expression and apparent race of these face stimuli were likely construed unambiguously. In the face categorization task in Experiment 2, emotion expression and race were both correctly classified on ≥95% of trials, supporting the assumption that both sources of information were clear and easy to identify (see Table S4).

primes were smiling and scowling Black men in one block of trials and smiling and scowling White men in the other block. Within a given block of trials, varying only one source of information (e.g., emotion expression) should render it more contextually distinctive, and thus more salient (Taylor & Fiske, 1978), than the other source of information (e.g., apparent race). Block order was counterbalanced and did not moderate racial bias.

Analysis Plan

Prior to all analyses, we excluded trials with RTs <100 ms and >1500 ms (Todd et al., 2021), which eliminated 2.4% of the data in both experiments. We also excluded error trials prior to RT analyses. Below, we report analyses pertinent to our focal hypotheses regarding information salience effects on racial bias. Full analysis of variance (ANOVA) tables appear in the Supplementary Materials (see Tables S1 and S2).

Results

Behavioral Analyses

Error Rates. A 2 (Salience) × 2 (Race Prime) × 2 (Expression Prime) × 2 (Target Object) ANOVA, with repeated measures on the last three factors, revealed a significant Salience × Race Prime × Target Object interaction, $F(1, 288) = 12.61$, $p < .001$, $\eta_p^2 = .04$, $CI_{90\%}$ [.01, .09], which did not vary by Expression Prime, $F(1, 288) = 0.04$, $p = .832$, $\eta_p^2 < .01$. To decompose this interaction, we examined the Race Prime × Target Object interaction separately per salience condition. (Table 2 displays descriptive statistics for the behavioral analyses in both experiments.)

The Race Prime × Target Object interaction was significant when race was salient, $F(1, 146) = 6.67$, $p = .011$, $\eta_p^2 = .04$, $CI_{90\%}$ [.01, .11]. Guns were misidentified as tools less often, $t(146) = -2.68$, $p = .008$, $d = -0.22$, $CI_{95\%}$ [-0.38, -0.06], whereas the misidentification of tools as guns did not significantly differ, $t(146) = 1.03$, $p = .305$, $d = 0.08$, $CI_{95\%}$ [-0.08, 0.25], after Black versus White primes. The Race Prime × Target Object interaction was also significant when emotion was salient, $F(1, 142) = 5.94$, $p = .016$, $\eta_p^2 = .04$, $CI_{90\%}$ [.004, .11], but the underlying pattern differed from that observed when race was salient. Whereas the misidentification of guns as tools did not significantly differ, $t(142) = -1.92$, $p = .057$, $d = -0.16$, $CI_{95\%}$ [-0.33, 0.00], tools were misidentified as guns *less* often, $t(142) = -4.58$, $p < .001$, $d = -0.38$, $CI_{95\%}$ [-0.55, -0.21], after Black versus White primes.

Table 2

Mean Error Rates and Correct Response Times by Condition (Experiments 1 and 2).

	Information salience and race prime			
	Race salient		Emotion salient	
Dependent Variable	Black prime	White prime	Black prime	White prime
Experiment 1				
Error rate (%)				
Gun	8.2 (27.5)	9.5 (29.3)	9.2 (28.9)	9.9 (29.9)
Tool	7.9 (27.1)	7.6 (26.5)	8.6 (28.1)	10.7 (30.9)
Correct RT (ms)				
Gun	287 (101)	291 (102)	289 (105)	281 (106)
Tool	321 (105)	314 (101)	317 (103)	317 (111)
Experiment 2				
Error rate (%)				
Gun	8.4 (27.7)	12.8 (33.4)	9.4 (29.3)	10.9 (31.1)
Tool	10.1 (30.2)	7.9 (26.9)	8.9 (28.6)	8.4 (27.8)
Correct RT (ms)				
Gun	287 (107)	301 (110)	283 (104)	287 (111)
Tool	332 (111)	319 (112)	317 (106)	311 (103)

Note. Values in parentheses are standard deviations. RT = response time.

Correct RTs. An identical ANOVA revealed a significant Salience × Race Prime × Target Object interaction, $F(1, 286) = 23.51$, $p < .001$, $\eta_p^2 = .08$, $CI_{90\%}$ [.03, .13], which did not vary by Expression Prime, $F(1, 286) < 0.01$, $p = .986$, $\eta_p^2 < .01$. We decomposed this interaction by examining the Race Prime × Target Object interaction separately per salience condition.

The Race Prime × Target Object interaction was significant when race was salient, $F(1, 146) = 20.42$, $p < .001$, $\eta_p^2 = .12$, $CI_{90\%}$ [.05, .21]. Gun identifications were faster, $t(146) = -2.62$, $p = .009$, $d = -0.22$, $CI_{95\%}$ [-0.38, -0.05], whereas tool identifications were slower, $t(146) = 4.85$, $p < .001$, $d = 0.40$, $CI_{95\%}$ [0.23, 0.57], after Black versus White primes. As with the error rates, the Race Prime × Target Object interaction was significant when emotion was salient, $F(1, 140) = 5.81$, $p = .017$, $\eta_p^2 = .04$, $CI_{90\%}$ [.004, .11], but the underlying pattern of racial bias differed. Gun identifications were *slower*, $t(141) = 2.53$, $p = .012$, $d = 0.21$, $CI_{95\%}$ [0.05, 0.38], whereas tool identifications did not significantly differ, $t(141) = -0.44$, $p = .662$, $d = -0.04$, $CI_{95\%}$ [-0.20, 0.13], after Black versus White primes.

Process Analyses

Next, we analyzed the DDM parameters by examining the 95% highest density interval ($HDI_{95\%}$) of the difference between posterior distributions of each parameter across relevant conditions. Differences with $HDI_{95\%}$ excluding 0 are considered credible. For each analysis, we report the most credible estimate of the raw difference, a Cohen's *d*, and the $HDI_{95\%}$ around *d*.[4] (Figure 2 displays the relative start point parameter estimates in both experiments; Figures S1–S3 display all other parameter estimates.)

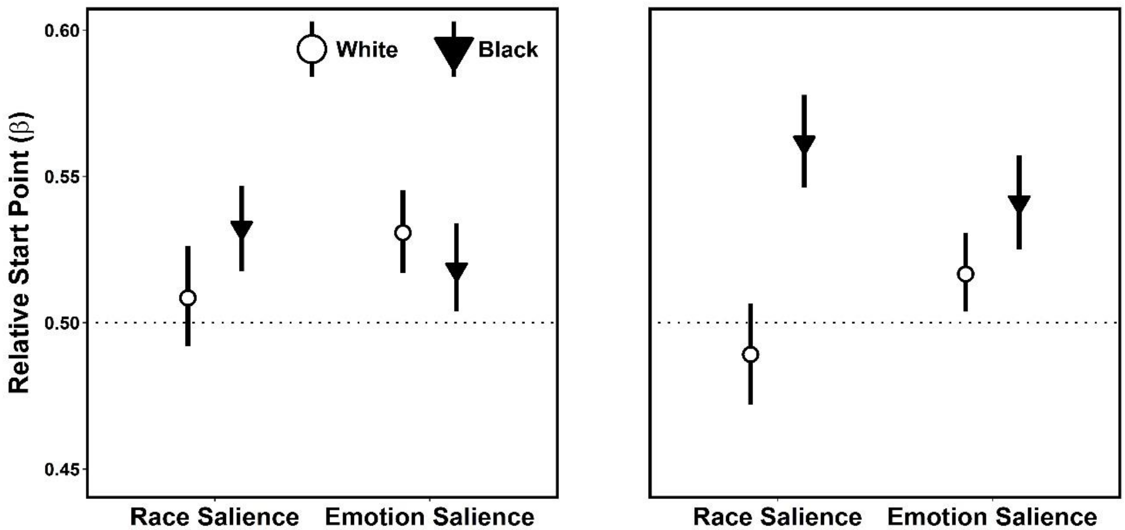
A Salience × Race Prime contrast on the relative start point (β) was credible, $\mu_{diff} = 0.02$, $d = 0.25$, $HDI_{95\%}$ [0.11, 0.41]. When race was salient, the decision process began closer to "gun" after Black versus White primes, $\mu_{diff} = -0.02$, $d = -0.33$, $HDI_{95\%}$ [-0.54, -0.13]. When emotion was salient, no credible racial bias emerged, $\mu_{diff} = 0.01$, $d = 0.18$, $HDI_{95\%}$ [-0.03, 0.38]. These

[4] For in-depth details regarding the DDM parameterization, estimation method, and analysis of model parameters, see Pleskac et al. (2018) and Todd et al. (2021).

findings align with an initial bias account: Salience-driven variation in racially biased starting positions in the decision process explain salience-driven moderation of racially biased behavior.

Figure 2

Relative Start Point (β) Parameter Estimates by Race Prime and Information Salience in Experiments 1 and 2



Notes. Markers reflect mean posterior predictions for White or Black prime trials; bars reflect 95% highest density intervals ($HDI_{95\%}$). The left plot displays estimates from Experiment 1; the right plot displays estimates from Experiment 2. The x-axis displays information salience condition (emotion, race).

A small but credible race prime effect emerged on the drift rate (δ), μ_{diff} = -0.14, *d* = -0.15, $HDI_{95\%}$ [-0.25, -0.06], but this effect did not vary by information salience, μ_{diff} = -0.05, *d* = -0.06, $HDI_{95\%}$ [-0.16, 0.03], or target object, μ_{diff} = 0.04, *d* = 0.04, $HDI_{95\%}$ [-0.06, 0.14]. Accumulated evidence from target objects was stronger following Black versus White primes, regardless of whether emotion or race information was more salient.

The race prime effect on threshold separation (α) was not credible, μ_{diff} = -0.02, *d* = -0.11, $HDI_{95\%}$ [-0.25, 0.02]. Finally, a small but credible race prime effect emerged on non-decision time (τ), μ_{diff} = -0.004, *d* = -0.10, $HDI_{95\%}$ [-0.20, -0.03], but this effect did not vary by information salience, μ_{diff} = -0.001, *d* = -0.02, $HDI_{95\%}$ [-0.12, 0.05], or target object, μ_{diff} = 0.00, *d* = 0.01, $HDI_{95\%}$ [-0.08, 0.09].

Discussion

In Experiment 1, racial bias was weaker when emotion versus race was salient. Process analyses failed to support an evidence accumulation account of this effect. Whereas stronger evidence was accumulated following Black versus White primes, neither target object nor information salience moderated this effect. Rather, process analyses supported an initial bias account: When race was salient, the weapon identification process began closer to the "gun" response after Black versus White primes. When emotion was salient, no credible starting point bias emerged. Descriptively, however, starting positions in the emotion-salient condition were *farther* from the "gun" response following Black versus White primes, mirroring the atypical pattern of behavior (e.g., *fewer* misidentified tools following Black versus White primes). Whether behavior assimilates toward (e.g., typical racial bias) or contrasts from (e.g., atypical racial bias) race stereotypes can vary by context (Bless & Schwarz, 2010), which raises questions about whether the atypical pattern in Experiment 1 stems from the contexts created by our blocking design. Experiment 2, therefore, sought to replicate these results using a different manipulation of information salience.

Experiment 2

Method

Participants

Prior work using a similar design (Todd et al., 2021, Experiment 1) revealed a large effect of information salience on racial bias in the WIT (Salience × Race Prime × Target Object interaction: η_p^2 = .139); however, because smaller effects are of theoretical interest, we set a target sample size (N = 258) affording ≥80% power to detect a η_p^2 = .03 (Faul et al., 2007). In total, 278 undergraduates consented to participate for course credit. We decided a priori to exclude data from participants who performed at or below chance (errors on ≥50% of trials) on the face categorization task (n = 1) or on any trial type in the WIT (n = 20). We also excluded data from one participant for whom a computer error caused the WIT to abort early. The final sample comprised 256 participants (73.4% women, 24.2% men, 1.2% non-binary; 12.7% White, 1.9% Black, 61.3% Asian, 15.2% Latinx, 4.7% multiracial; M_{age} = 19.4, SD = 2.0).

Procedure

Participants first completed a face categorization task (Todd et al., 2021) wherein they viewed one of two stimulus sets of facial images, each containing a randomly selected batch of 24 of the 48 facial images from Experiment 1. Both stimulus sets contained equal numbers of male faces varying in apparent race and posed emotion expression. Depending on information salience condition, participants were randomly assigned to classify the faces by *race* (Black vs. White) or by *emotion expression* (angry vs. happy) via key press. The images appeared one-by-one and remained on screen until participants responded, for a total of 72 trials.

Next, participants completed a WIT that deviated from the WIT in Experiment 1 in two ways. First, the face primes were the other set of 24 facial images not used during the face categorization task. We counterbalanced which stimulus set was used for the face categorization task and the WIT. Using different facial stimuli in the two tasks allowed us to rule out an event coding account (Hommel et al., 2001) whereby memory of specific responses toward specific faces in the face categorization task might affect responses toward those same faces in the WIT. Second, the face prime × target object combinations were fully integrated within a single block of 288 experimental trials that were preceded by 12 practice trials.

Results

Behavioral Analyses

Error Rates. A 2 (Salience) × 2 (Race Prime) × 2 (Expression Prime) × 2 (Target Object) ANOVA, with repeated measures on the last three factors, revealed a significant Salience × Race Prime × Target Object interaction, $F(1, 253) = 17.65$, $p < .001$, $\eta_p^2 = .07$, $CI_{90\%}$ [.02, .12], which did not vary by Expression Prime, $F(1, 253) = 1.39$, $p = .239$, $\eta_p^2 < .01$, $CI_{90\%}$ [.00, .03]. We decomposed this interaction by examining the Race Prime × Target Object separately per salience condition.

The Race Prime × Target Object interaction was significant when race was salient, $F(1, 127) = 53.45$, $p < .001$, $\eta_p^2 = .29$, $CI_{90\%}$ [.19, .40]. Guns were misidentified as tools less often, $t(127) = -7.82$, $p < .001$, $d = -0.69$, $CI_{95\%}$ [-0.88, -0.50], whereas tools were misidentified as guns more often, $t(127) = 4.21$, $p < .001$, $d = 0.37$, $CI_{95\%}$ [0.19, 0.55], after Black versus White primes. The Race Prime × Target Object interaction was also significant, though much smaller, when emotion was salient, $F(1, 126) = 9.51$, $p = .003$, $\eta_p^2 = .07$, $CI_{90\%}$ [.02, .15]. Unlike Experiment 1, a more typical pattern of racial bias emerged here: Guns were misidentified as tools less often, $t(126) = -3.09$, $p = .002$, $d = -0.27$, $CI_{95\%}$ [-0.45, -0.10], whereas the misidentification of tools as guns did not significantly differ, $t(126) = 1.35$, $p = .178$, $d = 0.12$, $CI_{95\%}$ [-0.05, 0.29], after Black versus White primes.

Correct RTs. An identical ANOVA revealed a significant Salience × Race Prime × Target Object interaction, $F(1, 253) = 14.00$, $p < .001$, $\eta_p^2 = .05$, $CI_{90\%}$ [.02, .10], which did not vary by Expression Prime, $F(1, 253) = 0.25$, $p = .621$, $\eta_p^2 < .01$. We decomposed this interaction by examining the Race Prime × Target Object interaction separately per salience condition.

The Race Prime × Target Object interaction was significant when race was salient, $F(1, 127) = 72.87$, $p < .001$, $\eta_p^2 = .37$, $CI_{90\%}$ [.26, .46]. Gun identifications were faster, $t(127) = -6.91$, $p < .001$, $d = -0.61$, $CI_{95\%}$ [-0.80, -0.42], whereas tool identifications were slower, $t(127) = 6.19$, $p < .001$, $d = 0.55$, $CI_{95\%}$ [0.36, 0.73], after Black versus White primes. The Race Prime × Target Object interaction was significant but smaller when emotion was salient, $F(1, 126) = 25.00$, $p < .001$, $\eta_p^2 = .17$, $CI_{90\%}$ [.08, .26]. Once again, unlike Experiment 1, a typical pattern of racial bias emerged here: Gun identifications were faster, $t(126) = -2.85$, $p = .005$, $d = -0.25$, $CI_{95\%}$ [-0.43, -0.08], whereas tool identifications were slower, $t(126) = 3.79$, $p < .001$, $d = 0.34$, $CI_{95\%}$ [0.16, 0.52], after Black versus White primes.

Process Analyses

A Salience × Race Prime contrast on the relative start point (β) was credible, $\mu_{diff} = 0.02$, $d = 0.40$, $HDI_{95\%}$ [0.25, 0.57]. When race was salient, the decision process began closer to "gun" after Black versus White primes, $\mu_{diff} = -0.07$, $d = -1.17$, $HDI_{95\%}$ [-1.45, -0.94]. Although starting point bias also emerged when emotion was salient, the effect was weaker, $\mu_{diff} = -0.02$, $d = -0.39$, $HDI_{95\%}$ [-0.60, -0.16]. As in Experiment 1, these findings align with an initial bias account.

A small but credible race prime effect emerged on the drift rate (δ), $\mu_{diff} = -0.13$, $d = -0.17$, $HDI_{95\%}$ [-0.27, -0.06], but this effect did not vary by information salience, $\mu_{diff} = 0.07$, $d = 0.10$, $HDI_{95\%}$ [-0.01, 0.20], or target object, $\mu_{diff} = 0.07$, $d = 0.10$, $HDI_{95\%}$ [-0.02, 0.20]. Stronger evidence was accumulated for the target objects following Black versus White primes, regardless of whether emotion or race information was more salient.

A small but credible race prime effect also emerged on threshold separation (α), $\mu_{diff} = -0.04$, $d = -0.24$, $HDI_{95\%}$ [-0.38, -0.09], but this effect did not vary by salience, $\mu_{diff} = 0.01$, $d = 0.07$, $HDI_{95\%}$ [-0.07, 0.22]. The amount of evidence required before responding was greater following Black versus White primes, regardless of whether emotion or race information was more salient. No credible effects emerged on non-decision time (τ).

Discussion

In Experiment 2, facial information salience again moderated racial bias in behavior, and these results again were better explained by an initial bias account than by an evidence accumulation account. The weapon identification decision process began closer to the "gun" response following Black versus White primes, but less so when emotion versus race information was more salient. Once again, stronger evidence accumulation following Black versus White primes did not vary by target object or which information was more salient.

General Discussion

Two experiments examined if and how the salience of facial information shapes racially biased weapon identification. Racial bias in behavior was weaker when the salience of emotion versus race was heightened, either by contextually augmenting the distinctiveness of emotion or race information during the WIT (Experiment 1) or by giving participants experience classifying faces by emotion expression or race prior to the WIT (Experiment 2). These results complement prior work suggesting that the salience of age versus race information can alter racially biased weapon identification (Jones & Fazio, 2010; Todd et al., 2021; see also Gawronski et al., 2010). Extending this prior work, we find that attending to comparatively less static and more affective-laden information communicated by facial expressions of emotion likewise can moderate racial bias.

Process analyses using diffusion modeling tested possible mechanisms by which facial information salience shapes racial bias. Both experiments supported an initial bias account, which posits that the decision process begins closer to the "gun" response on encountering a

Black versus White face prime, and that this starting point bias is weaker when attending to information besides race than when attending to race. Indeed, racially biased starting positions were either eliminated (Experiment 1) or weaker (Experiment 2) when emotion versus race was salient. Considered alongside previous findings of similar moderation by age salience (Todd et al., 2021), these results support the initial bias account as a mechanism whereby attending to person information besides race lowers the likelihood of favoring the “gun” response before the object’s appearance, relative to attending to race-related information.

Future research should test the generalizability of the initial bias account across other sources of salient facial information and different social groups. For example, information salience also shapes gender-stereotypic threat impressions (Rees et al., 2022), but it remains unclear where in the decision process these effects emerge. In addition, because we used only male face primes, future research should test whether racially biased weapon identification evoked by Black versus White *women* (Thiem et al., 2019) is likewise shaped by informational salience (cf. Petsko et al., 2022) and, if so, whether it is best explained by an initial bias account.

Notably, neither of our experiments replicated prior findings that the mere availability of emotion information in the face primes moderates racially biased weapon identification. Whereas Kubota and Ito (2014) found that racial bias emerged when face primes were scowling but not when they were smiling (see also Raissi & Steele, 2021), here emotion expressions in the face primes failed to moderate racial bias (despite these emotion expressions being easily detected; see footnote 3 and Table S4). Future work should determine whether the mere availability of emotion information in a target person’s face or the further heightening of its salience is required to moderate weapon-related racial bias.

Race information commonly detectable in faces can activate stereotypes (e.g., Black–gun association), which may bias decisions in potentially life-altering ways (e.g., misidentifying innocuous objects as guns; Payne & Correll, 2020). Our findings indicate that attending to emotion versus race information can weaken racial bias in weapon identification. Furthermore, this phenomenon can be explained by salience-driven changes at the start of the decision process. Racial biases favoring a “gun” response before the object’s onset were weaker when emotion versus race was more salient, pointing to a mechanism whereby the salience of person information moderates racially biased decision-making.

References

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, *15*, 88-93.

Bless, H., & Schwarz, N. (2010). Mental construal and the emergence of assimilation and contrast effects: The inclusion/exclusion model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 42, pp. 319-373). San Diego, CA: Elsevier Academic Press.

Correll, J., Wittenbrink, B., Crawford, M. T., & Sadler, M. S. (2015). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology, 108*, 219-233.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorisation influence spontaneous evaluations of multiply categorisable objects? *Cognition and Emotion*, *24*, 1008-1025.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133-168). New York: Guilford Press.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849-878.

Johnson, D. J., Cesario, J., & Pleskac, T. J. (2018). How prior information and police experience impact decisions to shoot. *Journal of Personality and Social Psychology, 115*, 601-623.

Jones, C. R., & Fazio, R. H. (2010). Person categorization and automatic racial stereotyping effects on weapon identification. *Personality and Social Psychology Bulletin, 36*, 1073-1085.

Kubota, J. T., & Ito, T. A. (2014). The role of expression and race in weapons identification. *Emotion, 14*, 1115-1124.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122-1135.

Macrae, C. N., & Cloutier, J. (2009). A matter of design: Priming context and person perception. *Journal of Experimental Social Psychology*, *45*(4), 1012-1015.

Niedenthal, P. M., & Brauer, M. (2012). Social functionality of human emotion. *Annual Review of Psychology*, *63*, 259-285.

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*, 181-192.

Payne, B. K., & Correll, J. (2020). Race, weapons, and the perception of threat. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 62, pp. 1-50). Academic Press.

Petsko, C. D., Rosette, A. S., & Bodenhausen, G. V. (2022). Through the looking glass: A lens-based account of intersectional stereotyping. *Journal of Personality and Social Psychology, 123*, 763-787

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

Raissi, A., & Steele, J. R. (2021). Does emotional expression moderate implicit racial bias? Examining bias following smiling and angry primes. *Social Cognition, 39*, 570-590.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260-281.

Rees, H. R., Sherman, J. W., Klauer, K. C., & Todd, A. R. (2022). On the use of gender categories and emotion categories in threat-based person impressions. *European Journal of Social Psychology, 52*, 597-610.

Richeson, J. A., & Trawalter, S. (2008). The threat of appearing prejudiced and race-based attentional biases. *Psychological Science*, *19*(2), 98-102.

Rivers, A. M. (2017). The weapons identification task: Recommendations for adequately powered research. *PLOS One*, *12*(6), e0177857.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249-288). New York, NY: Academic Press.

Thiem, K. C., Neel, R., Simpson, A. J., & Todd, A. R. (2019). Are Black women and girls associated with danger? Implicit racial bias at the intersection of target age and gender. *Personality and Social Psychology Bulletin, 45*, 1427-1439.

Todd, A. R., Johnson, D. J., Lassetter, B., Neel, R., Simpson, A. J., & Cesario, J. (2021). Category salience and racial bias in weapon identification: A diffusion modeling approach. *Journal of Personality and Social Psychology, 120*, 672-693.

Todd, A. R., Thiem, K. C., & Neel, R. (2016). Does seeing faces of young Black boys facilitate the identification of threatening stimuli? *Psychological Science, 27*, 384-393.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12*, 455-460.