Within-person dynamics of attention to race and expression of race bias: a real-time test of the self-regulation of prejudice model

Paul J. Brancaleone [1], Roberto U. Cofresí [2], Hannah I. Volpert-Esmond[3], David M. Amodio[4], Tiffany A. Ito[5], Bruce D. Bartholow[1,*]

[1]Department of Psychological and Brain Sciences, University of Iowa, 340 Iowa Ave, Iowa City, IA 52242, United States
[2]Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, United States
[3]Department of Psychology, University of Texas at El Paso, 500 W University, El Paso, TX 79902, United States
[4]Department of Psychology, University of Amsterdam, Postbus 15900, 1001 NK, Amsterdam, The Netherlands
[5]Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO 80309-0345, USA
[*]Corresponding author. Department of Psychological and Brain Sciences, University of Iowa, 472 PBSB, Iowa City, IA 52242, USA.
E-mail: bruce-bartholow@uiowa.edu.

Abstract

Prior research has supported some aspects of a theorized prejudice self-regulation model. We provide the first test of the full model-based process of bias regulation as it unfolds in real time. Event-related potentials (ERPs) were recorded from White undergraduates at two large American universities ($N = 130$; 40% female) during a racial stereotype priming task. Attention to Black male face primes, indexed by the P2 ERP, increased following self-regulation failures. In turn, within-person, trial-to-trial variability in attention to Black male faces predicted variability in bias expression. The latter effect was moderated by individual differences in internal motivation to respond without prejudice (IMS). Specifically, among lower-IMS individuals, trials in which Black faces elicited relatively larger P2 amplitudes (relative to an individual's own average P2 amplitude) were associated with increased behavioral race bias. In contrast, and consistent with theory, among higher-IMS individuals trials in which Black faces elicited larger relative P2 amplitudes were associated with *decreased* bias. Findings provide direct evidence supporting the temporal sequencing of race-bias regulation and identify within-person variability in attention to race as a potential mechanism for determining when and in whom bias will be regulated.

Keywords: race bias; attention; social categorization; event-related potentials; cognitive control

Introduction

Race bias arises when a person is categorized as a member of a stigmatized racial group, thereby activating stereotypes that bias judgement and behavior (see Wheeler and Petty 2001). Variability in bias regulation often is ascribed to between-person differences in perceivers' ability (e.g. Ito et al. 2015; Payne, 2005) or motivation (e.g. Devine et al. 1991, 2002) to exert control. Yet, bias regulation also varies within persons (e.g. Plant and Devine 2003, 2009), a fact few models consider (see Connor and Evers, 2020). Here, we propose and test a dynamic extension of the Self-regulation of Prejudice (SRP) model (Monteith 1993, Monteith et al. 2002) in which anti-Black bias is predicted to vary as a function of both between-person differences in internal motivation to respond without prejudice and within-person fluctuations in attention to race-related cues.

Motivational conflict and attention to race

According to the SRP model (Monteith 1993), unintentional expression of race bias elicits an aversive state of motivational conflict among individuals who value nonprejudiced behavior, prompting them to be vigilant for cues, such as race, signaling that bias could surface (Monteith et al. 2002). Consistent with this idea, failure to regulate bias in laboratory tasks elicits a pronounced error-related negativity (ERN; see Amodio et al. 2004, 2008), a component of the event-related brain potential (ERP) reflecting the degree and aversiveness of conflict between intended and actual responses (Hajcak and Foti, 2008; Inzlicht, et al., 2015). This posited sequence is analogous to the post-error (e.g. Egner and Hirsch 2005; Laming 1979) and postconflict (e.g. Gratton et al. 1992; Larson, et al., 2012) attention enhancement proposed in domain-general models of self-regulation. Whether a

failure to regulate bias in one instance enhances attention to race in the next has never been directly examined. The present study provides the first real-time test of this idea.

Attention to race as antecedent to bias (and its regulation)

A vast literature indicates that attending to race induces racial categorization and/or stereotype activation, thereby facilitating bias (e.g. Kunda and Spencer 2003). Supporting this perspective, enhanced processing of Black (vs. White) faces prompts greater reliance on stereotypes (e.g. Correll et al. 2006; Ofan, et al., 2011), whereas drawing attention away from social categories reduces stereotyping (e.g. Ito and Tomelleri, 2017; Macrae, et al., 1997; Wheeler and Fiske 2005). By contrast, the SRP model posits that, among people with low-prejudice personal standards, race can provide a strategic cue to *inhibit* bias (Monteith 1993, Monteith et al. 2002, 2009). Thus, among such individuals, greater attention to race in a given instance should *reduce* bias.

No prior studies have directly linked attention to race with in-the-moment regulation of race bias. Testing this idea requires a conceptual framework that acknowledges the temporally varying nature of visual attention (see Esterman and Rothlein 2019) and a method capable of capturing its moment-to-moment variability. Measuring ERPs as individuals attempt to regulate race bias provides such a method. The face-elicited P2 ERP is highly sensitive to race. In particular, Black faces typically elicit larger P2s than White faces (see Ito and Senholzi 2013, Amodio and Cikara 2021), a pattern observed in both White and Black perceivers in the USA (Volpert-Esmond and Bartholow, 2019, Volpert-Esmond and Bartholow 2021; but see Dickter and Bartholow 2007), in Chinese perceivers (Zhou et al. 2020), and in White European and North American perceivers living in China (Zhang et al. 2023). Several lines of work suggest face-elicited P2 amplitude reflects the extent to which attention is allocated for distinguishing racial categories. Racially ambiguous faces elicit smaller P2s than faces that more clearly represent racial outgroups (Willadsen-Jensen and Ito 2006, 2008, Ito et al. 2015). Visual fixation between the eyes elicits larger P2 than does fixation on the forehead (Volpert-Esmond et al. 2017, Volpert-Esmond and Bartholow, 2019), an area of the face that conveys less racial category information (Bülthoff, et al., 2021; Hills and Lewis 2006). Finally, P2 amplitude elicited by a given face predicts the speed with which that face can be racially categorized (Volpert-Esmond and Bartholow 2021).

Given the often-tight coupling between race categorization and bias (see Kawakami et al. 2017; Macrae and Bodenhausen, 2000), these findings suggest that within-person variability in face-elicited P2 amplitude could predict within-person variability in bias. In-principle support for this idea is provided by studies reporting that average (between-person) face-elicited P2 amplitude corresponds with averaged bias (Correll et al. 2006, Amodio 2010, Amodio and Swencionis 2018). Existing studies are limited by an implicit assumption that both bias and attention to race are stable within individuals, making only between-person (or between-condition) variability of interest. Multilevel modeling of ERP data (see Volpert-Esmond et al. 2018, 2021) allows testing associations at the level of individual trials, shifting the focus from between-person averages to within-person, trial-by-trial variability (see Vaughan and Birney 2023). We applied this approach to interrogate whether within-person variability in attention to race predicts when bias is likely to be regulated, and whether the form of this association varies according to between-person differences in low-prejudice personal standards.

The present research

This study provides the first real-time test of the SRP model's predictions concerning effects of motivational conflict on attention to race and effects of attention to race on bias expression (also see Volpert-Esmond et al. 2025). Bias was indexed by performance during the Weapons Identification Task (WIT; Payne 2001), which characterizes bias as faster and more accurate classification of guns relative to tools following exposure to Black male (vs. White male) faces (see Payne and Correll 2020). Between-person differences in low-prejudice personal standards were indexed by responses on the internal motivation to respond without prejudice scale (IMS; Plant and Devine 1998). IMS is more theoretically relevant than external motivation (EMS) in this context given the SRP model's focus on White perceivers' personal standards.

We advanced the following predictions (see Fig. 1). First, we expected the P2 elicited by Black (but not White) male faces to increase following errors versus correct responses (H1a). Extending this prediction to a continuous index of motivational conflict, we hypothesized that a larger (within-person) ERN on a given error trial, relative to an individual's average ERN, would predict a larger Black face-elicited P2 on the next trial (H2a). We expected these effects to be larger among higher-relative to lower-IMS individuals (H1b & H2b). Finally, we predicted that face-elicited P2 amplitude on a given Black-face trial would be associated with the degree of bias in response time (RT; H3a) and accuracy (H4a) to classify guns relative to tools, and critically, that these effects would be positive (i.e. larger P2 leading to more bias) among lower-IMS individuals and negative (i.e. larger P2 leading to *less* bias) among higher-IMS
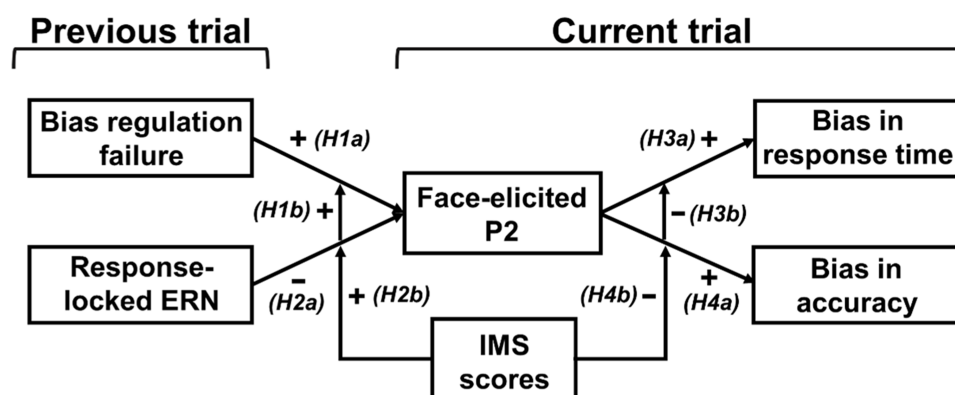
Figure 1. Conceptual model depicting predicted associations among key study variables.
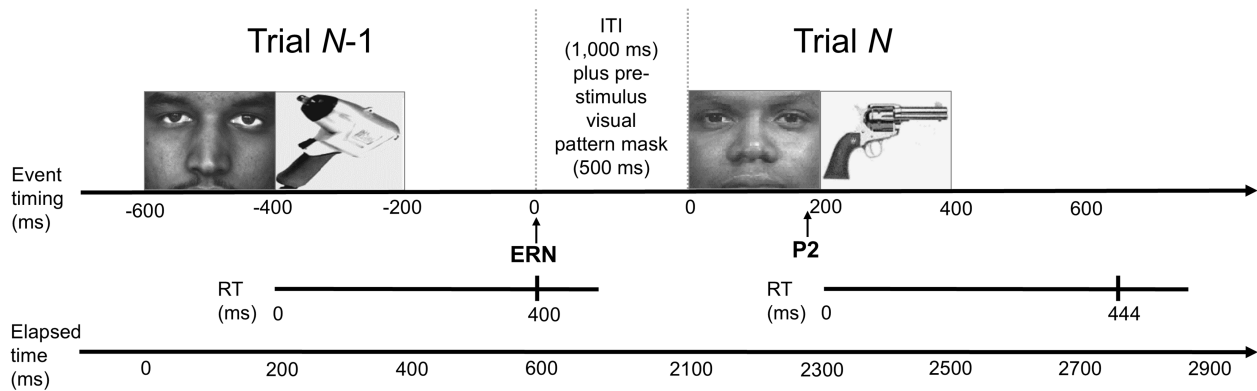
Figure 2. Timeline of a hypothetical two-trial sequence depicting the relative timing of the response-locked ERN on one trial and the face-elicited P2 on the following trial. (a) **Note.** For simplicity, the timeline does not include the pre- and post-stimuli visual pattern masks that were included on every trial. Vertical lines bisecting the reaction time (RT) timelines mark the mean error RT (Trial N-1) and correct RT (Trial N) time-locked to target onset.

individuals. Figure 2 provides a schematic of a hypothetical two-trial sequence depicting the temporal relationships among events of interest.

Method

Participants

Undergraduates ($N = 485$) were recruited from the University of Colorado Boulder and the University of Missouri to participate in a study on cognitive abilities. Individuals were excluded if they reported history of neurological disorder/disease, learning disability, or head trauma resulting in loss of consciousness for $\geq 2\,\text{min}$. A subset of participants ($n = 139$) completed the WIT while the electroencephalogram (EEG) was recorded. The present report includes data from White participants ($n = 130$) in this EEG subsample (60% male; $M_{age} = 19.84\,\text{years}$). Three participants were excluded due to missing EEG data or trigger codes and four others were excluded due to excessive EEG artifact, leaving a sample of 123 for data analysis. Behavioral data from the task used here for the full sample were reported in Ito et al. (2015). Response-locked ERPs were previously reported in Volpert-Esmond et al. (2018).

Materials

Weapons identification task

The weapons identification task (WIT; Payne 2001) is a speeded classification task in which cropped images of faces (of young White and Black men) precede presentation of target objects (tools and handguns) to be classified via button press. On each trial, a pattern mask (500 ms) preceded one of eight randomly selected face primes (grayscale; 16 total) (200 ms), followed immediately by a grayscale gun or tool target image (4 of each object; 200 ms), and finally, another visual pattern mask (300 ms). A 500 ms response deadline encompassed the duration of the target and post-target visual mask. Responses made after the deadline elicited a "Too Slow!" message (but were still recorded). Trials were separated by a 1 s ITI.

Motivation to respond without prejudice

Individual differences in motivations to respond without prejudice toward Black people were measured with the Internal and External Motivation to Respond Without Prejudice scales (Plant and Devine 1998). The 5-item internal motivation scale (IMS) assesses the degree to which people respond without prejudice

because it is personally important to them (e.g. "I am personally motivated by my beliefs to be nonprejudiced toward Black people"). The 5-item external motivation scale (EMS) assesses the degree to which people respond without prejudice to avoid social disapproval (e.g. "I try to act nonprejudiced toward Black people because of pressure from others"). Item scores were summed to create IMS ($M = 34.95$, s.d. $= 7.49$) and EMS ($M = 24.73$, s.d. $= 8.14$) scores, which were modestly correlated, $r(121) = 0.21$, $P = .018$. Internal consistencies for the total scale ($\alpha = 0.81$), the IMS ($\alpha = 0.81$), and the EMS ($\alpha = 0.78$), were acceptable.

EEG recording and ERP quantification

EEG was recorded from 32 tin electrodes in standard locations (10-10 system; Acharya et al. 2016), referenced to the right mastoid and re-referenced offline to an average of the mastoids. The EEG was sampled at 1000 Hz, filtered online at 0.10–40 Hz, and amplified with Neuroscan Synamps2 amplifiers (Compumedics, Charlotte, NC, USA). Impedances were kept below 10 KΩ. Eye movements were recorded and utilized offline for ocular artifact removal (Gratton et al. 1983). Epochs included 200 ms baseline periods and were derived prior to removing artifact-contaminated epochs (voltage deflections $\pm 75\,\mu V$) and to deriving individual trial responses and participant averages. For the ERN, response-locked data were down-sampled to 500 Hz, baseline correction was performed –600 to –400 ms pre-response, and epochs were further filtered at 1–15 Hz (96 db roll-off). Based on visual inspection of grand averages (Fig. 3) and on prior research (e.g. Volpert-Esmond et al. 2017), the face-elicited P2 was quantified as the mean amplitude 130-280 ms following face onset at central-parietal electrodes. ERN amplitude was quantified from the response-locked waveforms as the mean amplitude –25 to 130 ms post-response at frontal-central electrodes.

Analytic approach

Given our theoretical and methodological emphasis on White individuals' anti-Black bias, outcome measures for most models were derived only from Black-prime trials (i.e. trials on which anti-Black bias is possible). (Results of models using data from White-prime trials, reported in Supplementary Material, indicated no associations among constructs of interest.) This approach was informed by three main considerations. First, the WIT is designed to assess the stereotypical associations between young Black men and danger (Payne and Correll 2020); the task provides no analogous stereotypical association with young White men
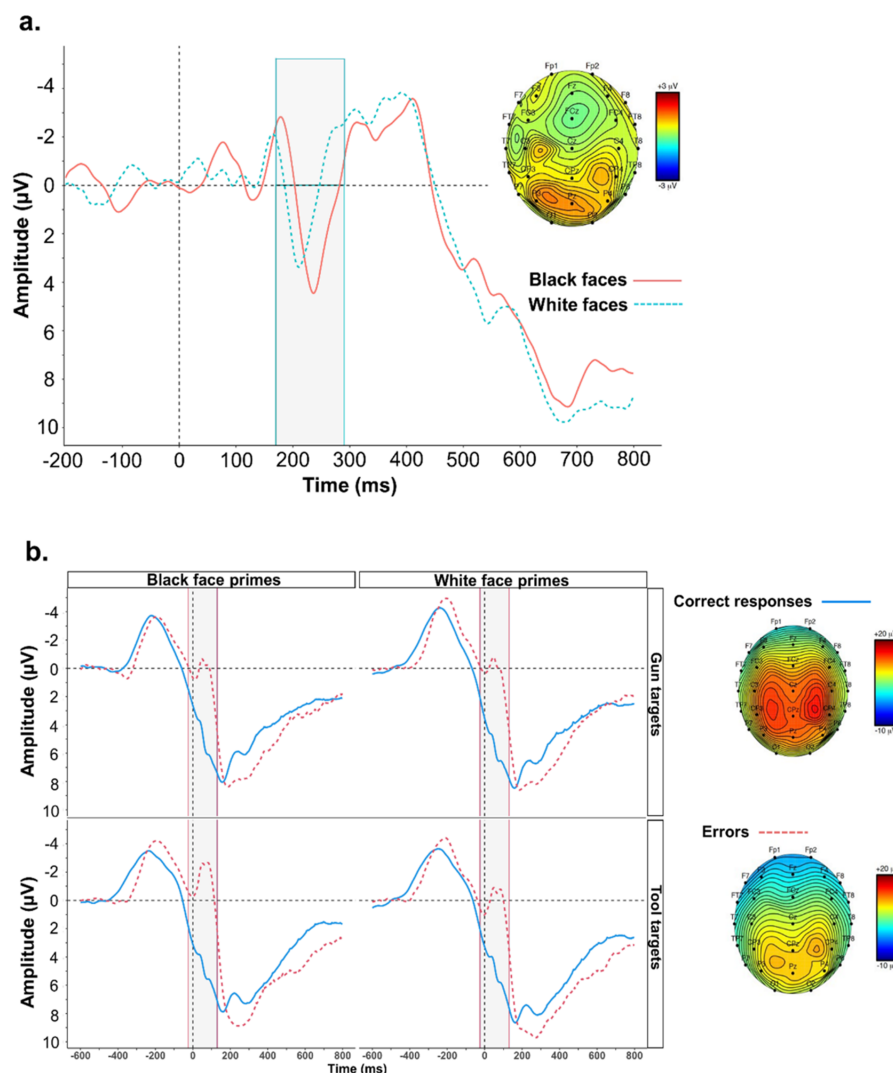
Figure 3. Grand average ERP waveforms and scalp topographies highlighting the face-elicited P2 and response-locked ERN responses. (a) **Note**. Shaded areas indicate the quantification windows for the P2 under electrode cites C3, C4, Cz, CPz, Pz, CP3, and CP4 (panel a) and ERN under electrodes F3, Fz, F4, FC3, FCz, and FC4 (b); scalp topography maps reflect voltage distributions during relevant quantification windows. Dashed vertical lines at 0 ms indicate the onset time of faces (a) and button press responses (b).

(Scherer and Lambert 2009). Second, the IMS and EMS focus exclusively on regulating responses toward Black people. Finally, dropping prime race as a predictor reduces the number of terms in the models, thereby simplifying statistical tests and focusing power—and control for Type I error—on hypothesis-relevant comparisons (Amodio et al. 2006).

Analyses used data from individual trials in the context of multilevel models (MLMs). MLMs permit derivation of model-estimated levels of a dependent variable on individual trials at levels of one or more predictors. Here, for example, the trial-level influence of P2 amplitude on accuracy bias was operationalized as the likelihood of accurately classifying a tool on a trial in which a (Black) face elicits P2 amplitude $= x$, relative to a trial in which the target is a gun and P2 amplitude $= x$. This approach permits testing whether within-person, trial-to-trial variability in a Level-1 predictor is meaningfully associated with variability in the outcome. Model specification for the random effects structure followed a data-driven approach by first specifying all possible random slopes for Level-1 predictors (i.e. the maximal model; Barr et al. 2013) and then systematically reducing random effects until (i) models converged

and (ii) AIC values indicated the best-fitting model. Continuous between-person predictors were grand-mean centered prior to analyses. Model-estimated slopes were generated for visualization and follow-up comparisons. Trials with RTs $\pm$ 3 s.d.'s from an individual's mean (~1.4% of trials) were removed prior to analyses.

Testing hypotheses regarding within-person, trial-level relationships between continuous predictor variables and relevant outcomes requires separating within-person and between-person contributions to predictors' variability. We applied a regression disaggregation procedure (Curran and Bauer, 2011) to separate ERP components into two predictors. The first predictor—each participant's mean P2 or ERN amplitude—was entered as a Level-2 (person-level) predictor and represents the between-person effect. The second predictor—P2 or ERN amplitude on each trial, centered around the participant's mean amplitude—was entered as a Level-1 (trial-level) predictor and represents the within-person effect. Including both predictors in the model permits interpretation of within-person effects on the dependent variable while holding constant any between-person differences in P2 or ERN amplitude.

According to Plant and Devine (1998), a White American's responses toward Black Americans may be guided primarily by internal motivation, by external motivation, by a combination of both, or by neither. This proposed configuration has been supported by latent profile analysis (Bamberg and Verkuyten 2022). Thus, although our predictions focused on differences related to internal motivation, EMS scores also were included in all primary models.

Transparency and openness

We report how we determined all data exclusions and all measures in the study. This article reports secondary analyses of existing data; hence, sample size was limited by the design of that prior study and was not based on any *a priori* power analysis. All data and analysis code and a description of the WIT stimuli and procedure are available at https://osf.io/dm5cw/. Other materials are available upon request. Data were analyzed using R, v.4.2.1 (R Core Team, 2022). Models were fit in the R packages lme4 v.1.1-35.3 (Bates et al. 2015) and lmerTest v.3.1-3 (Kuznetsova et al. 2017). To derive model-estimated slopes, we used the emmeans package, v.1.10.1 (Length et al. 2022) and plotted them using ggplot2, v.3.5.1 (Wickham 2016). Estimates of effect size (Semipartial R^2) were computed with r2glmm, v.0.1.2 (Jaeger et al. 2017). This study's design and analyses were not preregistered. We acknowledge this as a limitation of this study that should temper the interpretation of results and encourage future researchers working in this area to preregister their analyses and/or study designs.

Results

Preliminary analyses (see Supplementary Materials) confirmed that the patterns evident in the ERPs (Fig. 3) conformed with previous findings: Black faces elicited larger P2s than White faces, and the ERN was larger during errors indicative of race bias (Black-tool errors). Analyses of RT and accuracy data confirmed patterns of bias consistently observed in the WIT: classification of guns was faster and more accurate following Black face primes, whereas classification of tools was faster and more accurate following White face primes (see Payne and Correll 2020). Henceforth, we define bias as the discrepancy in the speed and accuracy with which tools are classified relative to guns on Black-prime trials (De Houwer 2019).

Primary analyses

*Sequential trial models predicting attention to race (H1 and H2)*

The first set of models tested whether attention to race (P2 amplitude) varies according to the accuracy of the response on the previous trial (H1a) and IMS levels (H1b). Face-elicited P2 amplitudes were submitted to a MLM including random slopes for race (Black, White). (Wilkinson notation: Current-trial P2 ~ Current-trial prime race*Previous-trial response*IMS*EMS+(1 + Current-trial prime race|Participant)). The model differed significantly from the random-intercept-only model, χ^2 (2) = 63.62, P < .001. Models including a random slope for previous response (correct, error) failed to converge and were not considered. Fixed effects are given in Table 1. (Significant effects from all primary analyses that are not directly relevant to our hypotheses are discussed in the Supplementary Material).

Main effects of race and previous response emerged, indicating that the P2 was larger when elicited by Black faces ($EMM = 2.12 \mu V$, $SE = 0.23$) compared to White faces ($EMM = 0.54 \mu V$, $SE = 0.24$) and following errors ($EMM = 1.49 \mu V$, $SE = 0.23$) compared to correct

Table 1. Fixed effects predicting face-elicited P2 amplitude from previous-trial response accuracy.

Fixed effect	b	SE	t-test	P	Semipartial R^2
Race	0.7964	0.066	12.0	<.001	0.00605
Previous response (PR)	–0.1326	0.048	–2.74	.006	0.00017
IMS	0.0012	0.031	0.04	.970	<0.00001
EMS	–0.0034	0.029	–0.12	.907	0.00001
Race × PR	**–0.0791**	**0.048**	**–1.65**	**.099**	**0.00006**
Race × IMS	0.0081	0.009	0.91	.366	0.00003
PR × IMS	0.0076	0.007	1.16	.245	0.00003
Race × EMS	0.0017	0.008	0.21	.834	<0.00001
PR × EMS	0.0135	0.006	2.29	.023	0.00012
IMS × EMS	–0.0031	0.003	–0.91	.365	0.00047
Race × PR × IMS	**–0.0023**	**0.006**	**–0.35**	**.723**	**<0.00001**
Race × PR × EMS	0.0035	0.006	0.59	.555	0.00001
Race × IMS × EMS	–0.0011	0.001	–1.12	.267	0.00006
PR × IMS × EMS	–0.0016	0.001	–2.39	.017	0.00013
Race × PR × IMS × EMS	0.0003	0.001	0.43	.700	<0.00001

Note: Model terms pertinent to hypotheses are presented in italicized boldface type. Race = race of the face eliciting the P2 (coded 1 = Black, –1 = White). PR = accuracy of the response on the previous trial (coded Correct = 1, Incorrect = -1). EMS = External motivation to respond without prejudice scale score; IMS = Internal motivation to respond without prejudice scale score.
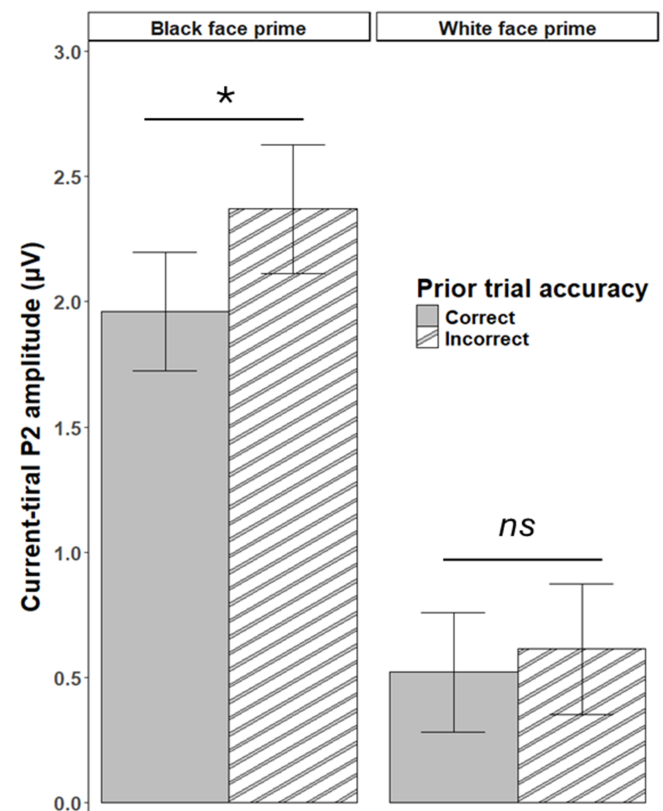
Figure 4. Face-elicited P2 amplitude as a function of target race and response accuracy on the previous trial.

responses ($EMM = 1.24 \mu V$, $SE = 0.24$). H1a is represented by the Race × Previous response interaction, which was not significant. However, given our directional, a priori prediction, we conducted follow-up simple effect contrasts to probe this effect (Fig. 4). Consistent with H1a, these contrasts indicated that whereas the P2 elicited by White faces was unaffected by previous-trial

Table 2. Fixed effects predicting face-elicited P2 amplitude from previous-trial ERN amplitude.

Fixed effect	b	SE	t-test	P	Semipartial R²
ERN (between)	−0.1456	0.0591	−2.47	.015	0.00585
ERN (within)	−0.0237	0.0094	−2.52	.012	0.00109
Race	0.9327	0.0996	9.36	<.001	0.01457
IMS	−0.0006	0.0318	−0.02	.985	<0.00001
EMS	−0.0064	0.0296	−0.22	.830	0.00005
ERN × Race	**0.0006**	**0.0093**	**0.06**	**.950**	**<0.00001**
ERN × IMS	0.0001	0.0013	0.08	.940	<0.00001
Race × IMS	−0.0037	0.0137	−0.27	.786	0.00001
ERN × EMS	0.0003	0.0012	0.23	.815	0.00001
Race × EMS	−0.0046	0.0121	−0.38	.707	0.00002
IMS × EMS	−0.0002	0.0035	−0.06	.948	<0.00001
ERN × Race × IMS	**−0.0013**	**0.0013**	**−1.02**	**.306**	**0.00018**
ERN × Race × EMS	0.0003	0.0012	0.26	.793	0.00001
ERN × IMS × EMS	0.0001	0.0001	0.92	.355	0.00015
Race × IMS × EMS	−0.0003	0.0014	−0.21	.833	0.00001
ERN × Race × IMS × EMS	−0.0001	0.0001	−0.51	.607	0.00004

Note: Model terms pertinent to hypotheses are presented in italicized boldface type. Race = race of the face eliciting the P2 (coded as 1 = Black, −1 = White). ERN (between) = between-person variance in (average) ERN amplitude; ERN (within) = within-person (trial-by-trial) variance in ERN amplitude. EMS = External motivation to respond without prejudice scale; IMS = Internal motivation to respond without prejudice scale.

Table 3. Fixed effects predicting trial-level reaction time from P2 amplitude.

Fixed effects	b	SE	t-test	P	Semipartial R²
P2 (between)	−2.0621	3.691	−1.63	.116	0.00392198
P2 (within)	−0.0279	1.266	−0.34	.677	0.00000686
IMS	0.2371	0.083	0.53	.619	0.00042359
EMS	−0.0967	0.450	−0.23	.813	0.00007723
Target	17.9994	0.419	16.70	<.001	0.04363304
P2 × IMS	−0.0040	1.078	−0.36	.667	0.00000760
P2 × EMS	0.0067	0.011	0.60	.567	0.00002133
IMS × EMS	0.0596	0.011	1.19	.225	0.00200074
P2 × Target	−0.0010	0.050	−0.01	.967	0.00000001
IMS × Target	−0.0534	0.083	−0.36	.763	0.00002150
EMS × Target	0.2670	0.147	1.99	.049	0.00058872
P2 × IMS x EMS	0.0003	0.138	0.22	.830	0.00000301
P2 × IMS x Target	**−0.0268**	**0.001**	**−2.40**	**.017**	**0.00034491**
P2 × EMS × Target	0.0138	0.011	1.23	.204	0.00009089
IMS × EMS × Target	0.0214	0.011	1.30	.225	0.00026088
P2 × IMS × EMS × Target	0.0016	0.017	1.19	.241	0.00008436

Note: Model terms pertinent to hypotheses are presented in italicized boldface type. Target was coded 1 = tool, −1 = gun. P2 (within) = within-person (trial-by-trial) variance in P2 amplitude; P2 (between) = between-person variance in P2 amplitude; IMS = internal motivation to respond without prejudice; EMS = external motivation to respond without prejudice. "P2" in all interaction terms refers to P2 (within). Black-prime trials only were included in this model.

accuracy ($b = -0.09$, $SE = 0.14$, 95% CI [−0.36, 0.18], $z = -0.68$, $P = .494$), the P2 elicited by Black faces was larger on trials following errors versus correct responses ($b = -0.41$, $SE = 0.14$, 95% CI [−0.68, −0.14], $z = -3.01$, $P = .003$). Support for H1a should be considered weak, however, given the nonsignificant P-value associated with the omnibus interaction. H1b posits that this interaction should be more pronounced for higher-IMS individuals; however, the Previous response × Race × IMS interaction was very small and not significant. Hence, H1b was not supported.

H2a predicts that the size of the ERN elicited by a given error will predict the magnitude of the P2 elicited by Black faces—but not White faces—on the next trial. To test this hypothesis, P2 amplitudes were submitted to a MLM with random intercepts by participant. (Current-trial P2 ~ Between-person ERN + Previous-trial within-person ERN*Prime*IMS*EMS+(1|Participant)). An alternative model including a random slope for race did not differ from the random-intercept-only model, χ^2 (2) = 1.114, $P = .573$. Fixed effects are given in Table 2. The model produced a main effect of within-person ERN amplitude, indicating that the ERN elicited on a given error trial predicted the magnitude of the face-elicited P2 on the following trial. However, this effect was not qualified by the predicted ERN × Race interaction (i.e. was not specific to Black faces; H2a), and the predicted ERN (within) x Prime race x IMS interaction (H2b) was not significant. Thus, H2a and H2b were not supported. Results of an alternative, more complex exploratory model examining the effects of the ERN elicited on specific trial types (e.g. previous-trial Black-tool ERN vs. White-tool ERN predicting current-trial P2) were largely null (see Supplementary Materials).

Within-trial models predicting bias (H3 and H4)

H3a states that larger P2 amplitude on a given Black-prime trial should correspond with greater bias in RT; H3b states that this pattern should be reversed among higher-IMS individuals (i.e. larger P2 facilitating greater bias *control*). To test this prediction, RT data were submitted to a MLM with random

intercepts by participant and random slopes of target type, (RT ~ P2 between-person + P2 within-person*Target*IMS*EMS+(1+ Target|Participant)) which differed significantly from the random-intercept-only model, χ^2 (2) = 90.49, $P < .001$. The maximal model failed to converge and was not considered. Fixed effects are given in Table 3.

The primary test of H3a is represented by the P2 (within) × Target interaction, which was very small and not significant. However, the P2 (within) × Target × IMS interaction—the primary test of H3b—was significant. To visualize this interaction, we derived model-estimated slopes of RT across the full range of P2 amplitudes present in the data (−40 to 48 μV), (Results did not change when the range of P2 amplitude values was restricted (e.g. −30 to 30 μV). Hence, we opted to retain the full range of P2 values present in the data) then plotted these values as a function of target type, separately for model-estimated "low-IMS" (−15) and "high-IMS" (+10) scores (±2 SD from the mean) (see Fig. 5). None of the simple slopes differed from zero (b's < 0.43, z's < 1.50), but the slope contrasts were significant for both low-IMS ($b = 0.077$, $SE = 0.38$, $z = 2.01$, $P = .044$, 95% CI [0.02, 1.52]) and high-IMS ($b = -0.56$, $SE = 0.27$, $z = -2.10$, $P = .036$, 95% CI [−1.09, −0.04]). Crucially, the slope patterns were consistent with the predictions under H3b. Among low-IMS individuals, relative increases in Black face-elicited P2 amplitude predicted RT differences indicative of increasing bias (faster classification of guns than tools). The pattern was reversed among higher-IMS individuals, such that RT differences indicative of bias *decreased* as a function of relative increases in Black face-elicited P2 amplitude.

H4a holds that a larger P2 on a given Black-prime trial should predict greater bias in accuracy (i.e. higher likelihood of correctly identifying a gun relative to a tool); H4b predicts the opposite pattern among higher-IMS individuals. To test these predictions, response accuracy on individual trials (0 = incorrect, 1 = correct) was submitted to a series of binomial generalized linear mixed-effects models. Probabilities were estimated and back-transformed from log odds using the emmeans package. The best-fitting model contained random intercepts by participant
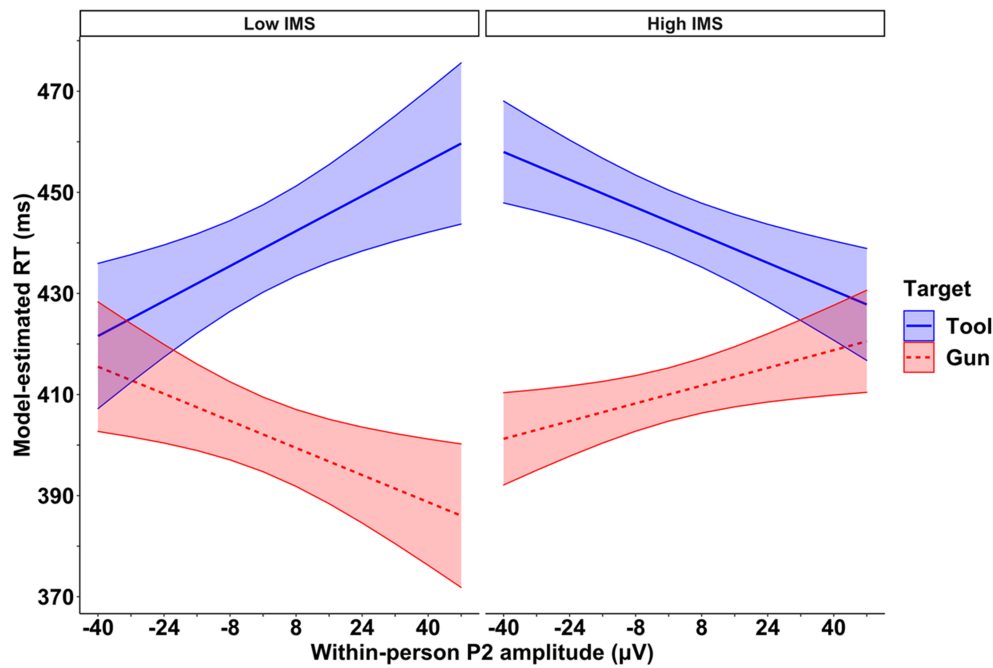
Figure 5. Latency to classify guns and tools as a function of (Black) face-elicited P2 amplitude and internal motivation to respond without prejudice.

Table 4. Fixed effects predicting trial-level probability of a correct response from P2 amplitude.

Fixed Effect	b	SE	z	P	Semipartial R^2
P2 (between)	−0.03688	0.02051	−1.80	.072	0.00130679
P2 (within)	−0.00005	0.00221	−0.02	.981	0.00000002
IMS	0.01122	0.00704	1.57	.117	0.00104362
EMS	−0.00135	0.00667	−0.20	.84	0.00001765
Target	−0.27415	0.03117	−8.80	<.001	0.01142687
P2 ×x IMS	0.00047	0.00029	1.60	.111	0.00011514
P2 × EMS	0.00003	0.00028	0.11	.913	0.00000048
IMS × EMS	0.00040	0.00070	−0.06	.957	0.00000141
P2 × Target	**0.00392**	**0.00221**	**1.80**	**.072**	**0.00015056**
Target × IMS	−0.00535	0.00421	−1.04	.205	0.00023719
Target × EMS	−0.01448	0.00391	−3.70	<.001	0.00196778
P2 × IMS × EMS	0.0000001	0.00001	−0.31	.760	0.00000396
P2 × IMS × Target	**0.00061**	**0.00029**	**2.08**	**.038**	**0.00019666**
P2 × EMS × Target	−0.00044	0.00028	−1.57	.116	0.00010793
IMS × EMS × Target	−0.00016	0.00046	0.35	.729	0.00001704
P2 × IMS × EMS × Target	−0.00006	0.00003	−1.81	.070	0.000013647

Note: Model terms pertinent to hypotheses are presented in italicized boldface type. Responses were coded as correct = 1, incorrect = 0. Target was coded as 1 = gun, −1 = tool. P2 (within) = within-person variability in P2 amplitude; P2 (between) = between-person variability in (average) P2 amplitude. IMS = internal motivation to respond without prejudice scale; EMS = external motivation to respond without prejudice scale. "P2" in all interaction terms refers to P2 (within). Black-prime trials only were included in this model. Unstandardized regression coefficients (*b*) represent log odds.

and random slopes of target, χ^2 (2) = 141.43, *P* < .001. (Accuracy ~ P2 between-person + P2 within-person*Target*IMS*EMS + (1 + Target|Participant)). Fixed effects are in Table 4.

The primary test of H4a is represented by the P2 (within) × Target interaction, which did not reach conventional statistical

significance. However, the P2 (within) × Target × IMS interaction—the test of H4b—was significant. To visualize this interaction, response accuracy was plotted as the (model-derived) probability of a correct response to tool targets and gun targets as a function of within-person P2 amplitudes, separately for model-derived "low-IMS" and "high-IMS" groups (see Fig. 6). Tukey-adjusted simple slope analyses indicated that, at lower levels of IMS, neither the slope for tool targets nor the slope for gun targets differed from zero (*b*'s < 0.018, zs < 1.64, *P*'s > .10) and the slopes did not differ from each another, *b* = −0.008, SE = 0.010, z = −0.815, *P* = .415, 95% CI [−0.03, 0.01]. At higher levels of IMS, however, the probability of correctly classifying a tool on a given Black-prime trial increased as a function of increasing face-elicited P2 amplitude, *b* = 0.014, SE = 0.005, z = 2.80, *P* = .005, 95% CI [0.004, 0.02]. Correct classification of guns was unrelated to face-elicited P2 amplitude (*b* = −0.005, SE = 0.01, z = −0.89, *P* = .376, 95% CI [−0.02, 0.01], but the slope contrast was significant, *b* = 0.018, SE = 0.01, z = 2.54, *P* = .011, 95% CI [0.004, 0.03]. Moreover, the slopes for tool targets differed significantly across the model-derived low-IMS and high-IMS groups, *b* = −0.025, SE = 0.001, z = −2.478, *P* = .013, 95% CI [−0.044, −0.005]. These patterns conform to the predictions under H4b.

Discussion

Despite decades of research identifying racial categorization as a key determinant of race bias (Hamilton 2015), and despite theory positing attention to race as a key to bias regulation (Monteith 1993), the present study is among the first to directly test the temporal dynamics of these factors as they unfold during attempts to regulate bias (also see Volpert-Esmond et al. 2025). Findings provided little support for the idea that control failure in one moment enhances attention to cues signaling the potential for bias in the next among individuals with low-prejudice personal standards. Post-hoc analyses revealed the P2 was larger on trials following errors, especially when accompanied by larger ERNs, but evidence for the specificity of this effect to Black
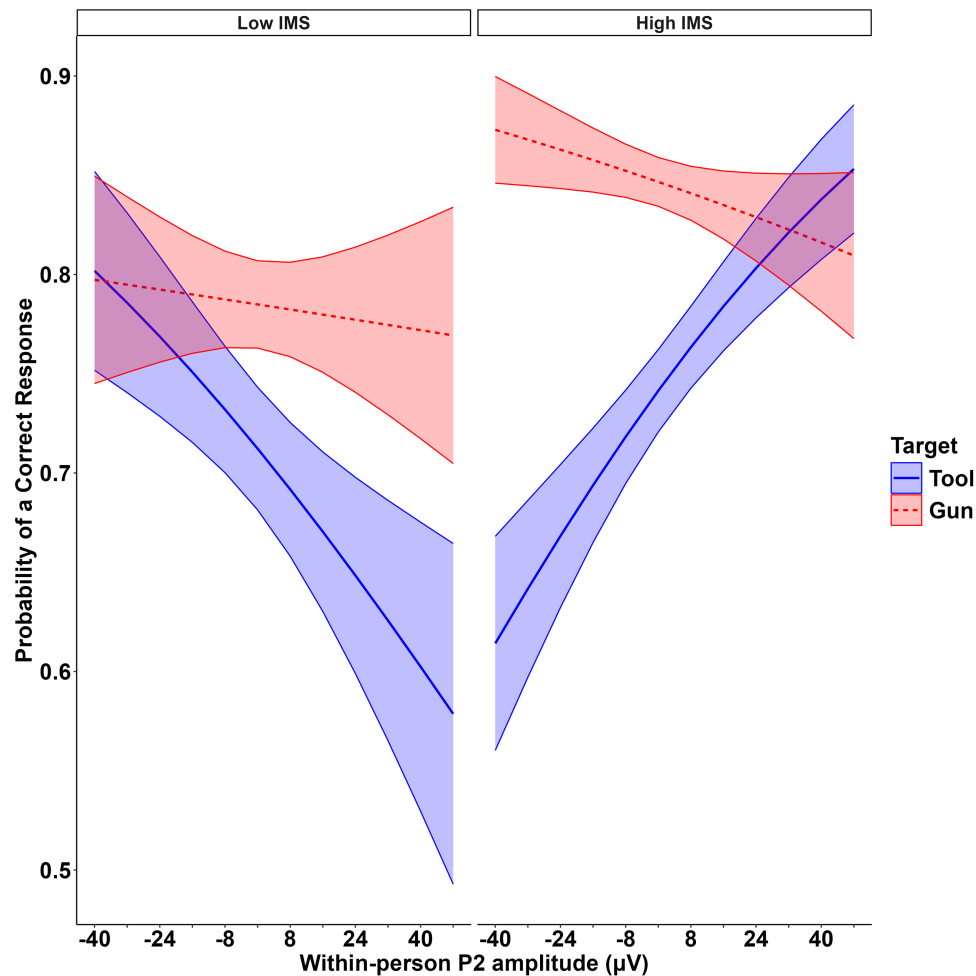
Figure 6. Model-estimated probability of correctly classifying tools and guns as a function of (Black) face-elicited P2 amplitude and internal motivation to respond without prejudice.

faces was weak (Fig. 4) and not limited to higher-IMS individuals. These findings are consistent with the idea that control failures prompt a largely nonspecific phasic increase in attention to task-related stimuli (Egner and Hirsch 2005, Walsh et al. 2011).

However, the data supported the prediction that variability in attention to race cues differentially affects bias depending upon perceivers' prejudice-related personal standards. Among lower-IMS individuals, within-person relative increases in P2 amplitude elicited by Black faces were accompanied by relative increases in bias (Figs 5 and 6). Critically, the opposite pattern emerged among higher-IMS participants, who expressed *less* bias during trials for which Black male faces elicited larger P2s. (Note the slopes in Figs 5 and 6 represent patterns in the model-estimated means in relative terms. That is, the data indicate that bias expression is more likely among high-IMS individuals when P2s elicited by Black faces are much smaller than an individual's own average, not when P2s are small or absent in absolute terms.) This finding, along with another recent report from our lab (Volpert-Esmond et al. 2025), provides the first direct support for the SRP model's prediction (Monteith 1993, Monteith et al. 2002) that individuals holding low-prejudice personal standards benefit from relative increases in attention to Black faces in terms of engaging control over bias. Using a different sample and a different bias task, one permitting a much longer decision window (4 s) than was used here, Volpert-Esmond et al. (2025) similarly found that the

magnitude of the P2 elicited by White men's and Black men's faces predicted race bias (in financial investments), and that the form of this association differed as a function of internalized motivation to respond without prejudice. These findings lend confidence that the pattern observed here is generalizable beyond the specific sample, task parameters, and index of bias used here.

Broader implications

These findings have implications beyond the SRP model. First, they contribute to models of "mental chronometry," the temporal dynamics of mental events (Coles et al. 1995), as applied to race categorization and its consequences (see Freeman and Ambady 2011). According to the continuous flow model of information processing (Eriksen and Schultz, 1979), stimulus evaluation unfolds over time as information about the stimulus is extracted and elaborated upon. Initial, partial information about the stimulus informs subsequent stages, including response output (Coles, et al., 1985). The more evaluative information is extracted early in processing, the stronger its influence on subsequent stages. This idea accords with Volpert-Esmond and Bartholow's (2021) finding that faces eliciting a larger relative P2 are categorized by race more quickly, and with Rees et al.'s (2020) finding that the strength of stereotype activation—and the difficulty controlling stereotype-based responding—depends on how deeply race-related information is processed. The present data extend this prior work, showing that the degree to which early attention is

directed to processing race during brief exposures to race-related information has consequences for regulating race bias.

The present findings also contribute to understanding the forms of control by which bias is regulated. Dual Mechanisms of Control theory (see Braver 2012) posits two forms of cognitive control that support goal-directed behavior. *Proactive control* pre-emptively supports intended responding via active maintenance of goal-relevant information in working memory; *reactive control* is recruited in-the-moment as a "late correction" mechanism when performance goals are threatened. The latter contributes to bias regulation when errors (Amodio et al. 2004; Bartholow, et al., 2012) or conflict (Bartholow et al. 2006) energize efforts to re-establish control. A role for *tonic proactive control* was demonstrated by Amodio and Swencionis (2018), who reported that increasing the proportion of stereotype-incongruent (i.e. Black-tool) trials in the WIT reduced both (average) face-elicited P2 amplitude and (average) behavioral bias. This approach mimics the idea first proposed by Gratton et al. (1992) that people use contextual information (i.e. the global level of interference in a situation) to strategically control their attention. But, when task conditions (e.g. equal proportions of high- and low-conflict trials) do not support use of such a strategy, self-regulation must depend on alternative mechanisms.

Amodio and Swencionis (2018) speculated that individuals with egalitarian beliefs (i.e. high-IMS) would rely on an interplay between what they dubbed *chronic proactive control*—trait-like vigilance for a situation or response rooted in chronic goals or beliefs—and *cued proactive control*—use of environmental cues signaling the potential for goal conflict to trigger retrieval of intentions from long-term memory (Ball and Brewer 2018, Bugg et al. 2013; Bugg and Smallwood, 2016). The present results are in-line with this idea. For high-IMS individuals, race cues appear to function as this kind of goal-relevant stimulus, helping to energize chronic intentions to be unbiased and/or to engage inhibitory control over stereotype-based responses. Future research could further these ideas by combining a manipulation of interference levels in the WIT (as in Amodio and Swencionis) with measurement of IMS, which could reveal strategic flexibility in preferred forms of control. We predict that, in low-interference blocks (i.e. when context does not trigger tonic proactive control), attention to race cues and their influence on behavioral bias will depend on levels of IMS, as in the present study. But, in high-interference blocks, all participants, regardless of IMS levels, should adopt a tonic proactive control strategy, evident in reduced attention to race cues.

Conclusions

In the USA, split-second decisions have too often determined whether young Black men are subjected to a form of bias that threatens their lives. Yet, the temporal dynamics of factors contributing to bias regulation remain little understood. While keeping in mind the study's limitations, including a predetermined sample size that constrained statistical power for testing interactions, the present findings advance understanding in this area by demonstrating, first and foremost, that mechanisms of bias are not stable within individuals but vary across time and interactions. Of greater importance, the findings show that a given individual is more likely to express bias when their early visual attention is allocated to distinguishing racial categories—unless that individual is motivated by nonprejudiced personal standards to control their bias. For those people, more strongly attending to race can mitigate bias. Thus, perhaps efforts should be directed

at bolstering internally guided motivations to control prejudiced responding, possibly through perspective-taking exercises (e.g. Lindsey et al. 2015).

Supplementary data

Supplementary data is available at *SCAN* online.

Funding

References

Acharya JN, Hani A, Cheek J *et al*. American Clinical Neurophysiology Society Guideline 2: Guidelines for standard electrode position nomenclature. *J Clin Neurophysiol* 2016;**33**:308–311. https://doi.org/10.1097/WNP.0000000000000316

Amodio DM. Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behavior. *J Cogn Neurosci* 2010;**22**:2609–17. https://doi.org/10.1162/jocn.2009.21395

Amodio DM, Cikara M. The social neuroscience of prejudice. *Annu Rev Psychol* 2021;**72**:439–69. https://doi.org/10.1146/annurev-psych-010419-050928

Amodio DM, Devine PG, Harmon-Jones E. Individual differences in the regulation of intergroup bias: the role of conflict monitoring and neural signals forcontrol. *J Pers Soc Psychol* 2008;**94**:60–74. https://doi.org/10.1037/0022-3514.94.1.60

Amodio DM, Harmon-Jones E, Devine PG *et al*. Neural signals for the detection of unintentional race bias. *Psychol Sci* 2004;**15**:88–93. https://doi.org/10.1111/j.0963-7214.2004.01502003.x

Amodio DM, Kubota JT, Harmon-Jones E *et al*. Alternative mechanisms for regulating racial responses according to internal vs external cues. *Soc Cogn Affect Neur* 2006;**1**:26–36. https://doi.org/10.1093/scan/nsl002

Amodio DM, Swencionis JK. Proactive control of implicit bias: a theoretical model and implications for behavior change. *J Pers Soc Psychol* 2018;**115**:255–75. https://doi.org/10.1037/pspi0000128

Ball BH, Brewer GA. Proactive control processes in event-based prospective memory: evidence from intraindividual variability and ex-Gaussian analyses. *J Exp Psychol* 2018;**44**:793–811. https://doi.org/10.1037/xlm0000489

Bamberg K, Verkuyten M. Internal and external motivation to respond without prejudice: a person-centered approach. *J Soc Psychol* 2022;**162**:435–54. https://doi.org/10.1080/00224545.2021.1917498

Barr DJ, Levy R, Scheepers C *et al*. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J Memory Lang* 2013;**68**:255–78.

Bartholow BD, Dickter CL, Sestir MA. Stereotype activation and control of race bias: cognitive control of inhibition and its impairment by alcohol. *J Pers Soc Psychol* 2006;**90**:272–87. https://doi.org/10.1037/0022-3514.90.2.272

Bartholow BD, Henry EA, Lust SA *et al*. Alcohol effects on performance monitoring and adjustment: Affect modulation and impairment of evaluative cognitive control. *J Abnorm Psychol* 2012;**121**:173–186. https://doi.org/10.1037/a0023664

Bates D, Machler M, Bolker B *et al*. Fitting linear mixed-effects mod-elsusing lme4. *J Stat Softw* 2015;**67**:1–48. https://doi.org/10.18637/jss.v067.i01

Braver TS. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn Sci* 2012;**16**:106–13. https://doi.org/10.1016/j.tics.2011.12.010

Bugg JM, McDaniel MA, Einstein GO. Event-based prospective remembering: an integration of prospective memory and cognitive control theories. In: Reisberg D (ed.), *The Oxford Handbook of Cognitive Psychology*. New York: Oxford University Press, 2013, 267–82.

Bugg JM, Smallwood A. The next trial will be conflicting! Effects of explicit congruency pre-cues on cognitive control. *Psychol Res* 2016;**80**:16–33. https://doi.org/10.1007/s00426-014-0638-5

Bülthoff I, Jung W, Armann RG *et al*. Predominance of eyes and surface information for face race categorization. *Sci Rep* 2021;**11**:1927. https://doi.org/10.1038/s41598-021-81476-1

Coles MGH, Gratton G, Bashore TR *et al*. A psychophysiological investigation of the continuous flow model of human information processing. *J Exp Psychol Human* 1985;**11**:529–553. https://doi.org/10.1037/0096-1523.11.5.529

Coles MGH, Smid HGOM, Scheffers MK *et al*. Mental chronometry and the study of human information processing. In: Rugg MD, Coles MGH (eds), *Electrophysiology of Mind: Event-related Brain Potentials and Cognition*. New York: Oxford University Press, 1995, 86–131.

Connor P, Evers ER. The bias of individuals (in Crowds): Why implicit bias is probably a noisily measured individual-level construct. *Perspect Psychol Sci* 2020;**15**:1329–1345. https://doi.org/10.1177/1745691620931492

Correll J, Urland GR, Ito TA. Event-related potentials and the decision to shoot: the role of threat perception and cognitive control. *J Exp Soc Psychol* 2006;**42**:120–28.

Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. *Annu Rev Psychol* 2011;**62**:583–619. https://doi.org/10.1146/annurev.psych.093008.100356

De Houwer J. Implicit bias is behavior: a functional-cognitive perspective on implicit bias. *Perspect Psychol Sci* 2019;**14**:835–40. https://doi.org/10.1177/1745691619855638

Devine PG, Monteith MJ, Zuwerink JR *et al*. Prejudice with and without compunction. *J Pers Soc Psychol* 1991;**60**:817–30. https://doi.org/10.1037/0022-3514.60.6.817

Devine PG, Plant EA, Amodio DM *et al*. The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice. *J Pers Soc Psychol* 2002;**82**:835–48. https://doi.org/10.1037/0022-3514.82.5.835

Dickter CL, Bartholow BD. Event-related brain potential evidence of ingroup and outgroup attention biases. *Soc Cogn Affect Neurosci* 2007;**2**:189–98. https://doi.org/10.1093/scan/nsm012

Egner T, Hirsch J. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat Neurosci* 2005;**8**:1784–90. https://doi.org/10.1038/nn1594

Eriksen CW, Schultz DW. Information processing in visual search: A continuous flow conception and experimental results. *Percept Psychophys* 1979;**25**:249–263. https://doi.org/10.3758/BF03198804

Esterman M, Rothlein D. Models of sustained attention. *Curr Opin Psychol* 2019;**29**:174–80. https://doi.org/10.1016/j.copsyc.2019.03.005

Freeman JB, Ambady N. A dynamic interactive model of person construal. *Psychol Rev* 2011;**118**:247–79. https://doi.org/10.1037/a0022327

Gratton G, Coles M, Donchin E. Optimizing the use of information: strategic control of activation of responses. *J Exp Psychol —General* 1992;**121**:480–506. https://doi.org/10.1037/0096-3445.121.4.480

Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* 1983;**55**:468–84. https://doi.org/10.1016/0013-4694(83)90135-9

Hajcak G, Foti D. Errors are aversive. *Psychol Sci* 2008;**19**:103–108. https://doi.org/10.1111/j.1467-9280.2008.02053.x

Hamilton DL () *Cognitive Processes in Stereotyping and Intergroup Behavior*. London, UK: Psychology Press, 2015.

Hills PJ, Lewis MB. Reducing the own-race bias in face recognition by shifting attention. *Q J Exp Psychol* 2006;**59**:996–1002. https://doi.org/10.1080/17470210600654750

Inzlicht M, Bartholow BD, Hirsh JB. Emotional foundations of cognitive control. *Trends in Cognitive Sciences* 2015;**19**:126–132. https://doi.org/10.1016/j.tics.2015.01.004

Ito TA, Friedman NP, Bartholow BD *et al*. Toward a comprehensive model of executive cognitive function in implicit racial bias. *J Pers Soc Psychol* 2015;**108**:187–218. https://doi.org/10.1037/a0038557

Ito TA, Senholzi KB. Us versus them: understanding the process of race perception with event-related brain potentials. *Vis Cogn* 2013;**21**:1096–120. https://doi.org/10.1080/13506285.2013.821430

Ito TA, Tomelleri S. Seeing is not stereotyping: the functional independence of categorization and stereotype activation. *Soc Cogn Affect Neurosci* 2017;**12**:758–64. https://doi.org/10.1093/scan/nsx009

Jaeger BC, Edwards LJ, Das K *et al*. An R^2 statistic foR fixed effects in the geneRalized lineaR mixed model. *J Appl Stat* 2017;**44**:1086–105. https://doi.org/10.1080/02664763.2016.1193725

Kawakami K, Amodio DM, Hugenberg K. Intergroup perception and cognition: an integrative framework for understanding the causes and consequences of social categorization. *Adv Exp Soc Psychol* 2017;**55**:1–80.

Kunda Z, Spencer SJ. When do stereotypes come to mind and when do they color judgment? *Psychol Bull* 2003;**129**:522–44. https://doi.org/10.1037/0033-2909.129.4.522

Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw* 2017;**82**:1–26. https://doi.org/10.18637/jss.v082.i13

Laming D. Choice reaction performance following an error. *Acta Psychologica* 1979;**43**:199–224. https://doi.org/10.1016/0001-6918(79)90026-X

Larson MJ, Clayson PE, Baldwin SA. Performance monitoring following conflict: Internal adjustments in cognitive control?. *Neuropsychologia* 2012;**50**:426–433. https://doi.org/10.1016/j.neuropsychologia.2011.12.021

Length RV, Buerkner P, Herve M *et al*. (2022). Emmeans R package. https://CRAN.R-project.org/package=emmeans (Accessed 10 June 2022)

Lindsey A, King E, Hebl M *et al*. The impact of method, motivation, and empathy on diversity training effectiveness. *J Bus Psychol* 2015;**30**:605–17. https://doi.org/10.1007/s10869-014-9384-3

Macrae CN, Bodenhausen GV. Social cognition: thinking categorically about others. *Annu Rev Psychol* 2000;**51**:93–120. https://doi.org/10.1146/annurev.psych.51.1.93

Macrae CN, Bodenhausen GV, Milne AB *et al*. On the activation of social stereotypes: The moderating role of processing objectives. *J Exp Soc Psychol* 1997;**33**:471–489. https://doi.org/10.1006/jesp.1997.1328

Monteith MJ. Self-regulation of prejudiced responses: implications for progress in prejudice reduction efforts. *J Pers Soc Psychol* 1993;**65**:469–85. https://doi.org/10.1037/0022-3514.65.3.469

Monteith MJ, Ashburn-Nardo L, Voils CI *et al*. Putting the brakes on prejudice: on the development and operation of cues for control. *J Pers Soc Psychol* 2002;**83**:1029–50. https://doi.org/10.1037/0022-3514.83.5.1029

Monteith MJ, Devine PG, Zuwerink JR. Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *J Pers Soc Psychol* 1993;**64**:198–210.

Monteith MJ, Lybarger JE, Woodcock A. Schooling the cognitive monster: the role of motivation in the regulation and control of prejudice. *Soc Pers Psychol Compass* 2009;**3**:211–26. https://doi.org/10.1111/j.1751-9004.2009.00177.x

Ofan RH, Rubin N Amodio DM Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing. *J Cognitive Neurosci* 2011;**23**:3153–3161. https://doi.org/10.1162/jocn_a_00014

Payne BK. Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *J Pers Soc Psychol* 2001;**81**:181–92. https://doi.org/10.1037/0022-3514.81.2.181

Payne BK. Conceptualizing control in social cognition: how executive functioning modulates the expression of automatic stereotyping. *J Pers Soc Psychol* 2005;**89**:488–503. https://doi.org/10.1037/0022-3514.89.4.488

Payne BK, Correll J. Race, weapons, and the perception of threat. *Adv Exp Soc Psychol* 2020;**62**:1–50.

Plant EA, Devine PG. Internal and external motivation to respond without prejudice. *J Pers Soc Psychol* 1998;**75**:811–32. https://doi.org/10.1037/0022-3514.75.3.811

Plant EA, Devine PG. The antecedents and implications of interracial anxiety. *Pers Soc Psychol Bull* 2003;**29**:790–801. https://doi.org/10.1177/0146167203029006011

Plant EA, Devine PG. The active control of prejudice: unpacking the intentions guiding control efforts. *J Pers Soc Psychol* 2009;**96**:640–52. https://doi.org/10.1037/a0012960

R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing* 2022. https://www.R-project.org

Rees HR, Ma DS, Sherman JW. Examining the relationships among categorization, stereotype activation, and stereotype application. *Pers Soc Psychol Bull* 2020;**46**:499–513. https://doi.org/10.1177/0146167219861431

Scherer LD, Lambert AJ. Contrast effects in priming paradigms: implications for theory and research on implicit attitudes. *J Pers Soc Psychol* 2009;**97**:383–403. https://doi.org/10.1037/a0015844

Vaughan AC, Birney DP. Within-individual variation in cognitive performance is not noise: why and how cognitive assessments should examine within-person performance. *J Intell* 2023;**11**:110. https://doi.org/10.3390/jintelligence11060110

Volpert-Esmond HI, Bartholow BD. Explicit categorization goals affect attention-related processing of race and gender during person construal. *J Exp Soc Psychol* 2019;**85**:103839. https://doi.org/10.1016/j.jesp.2019.103839

Volpert-Esmond HI, Bartholow BD. A functional coupling of brain and behavior during social categorization of faces. *Pers Soc Psychol Bull* 2021;**47**:1580–95. https://doi.org/10.1177/0146167220976688

Volpert-Esmond HI, Bray J, Levsen MP *et al*. Share the wealth: neurophysiological and motivational mechanisms contributing to racial discrimination in economic decision making. *J Exp Soc Psychol* 2025;**116**:104683. https://doi.org/10.1016/j.jesp.2024.104683

Volpert-Esmond HI, Merkle E, Bartholow BD. The iterative nature of person construal: Evidence from event-related potentials. *Soc Cogn Affect Neurosci* 2017;**12**:1097–107. https://doi.org/10.1093/scan/nsx048

Volpert-Esmond HI, Merkle EC, Levsen MP *et al*. Using trial-level data and multilevel modeling to investigate within-task change in event-related potentials. *Psychophysiology* 2018;**55**:e13044. https://doi.org/10.1111/psyp.13044

Volpert-Esmond HI, Page-Gould E, Bartholow BD. Using multilevel models for the analysis of event-related potentials. *Int J Psychophysiol* 2021;**162**:145–56. https://doi.org/10.1016/j.ijpsycho.2021.02.006

Walsh BJ, Buonocore MH, Carter CS *et al*. Integrating conflict detection and attentional control mechanisms. *J Cogn Neurosci* 2011;**23**:2211–21. https://doi.org/10.1162/jocn.2010.21595

Wheeler ME, Fiske ST. Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychol Sci* 2005;**16**:56–63. https://doi.org/10.1111/j.0956-7976.2005.00780.x

Wheeler SC, Petty RE. The effects of stereotype activation on behavior: a review of possible mechanisms. *Psychol Bull* 2001;**127**:797–826. https://doi.org/10.1037/0033-2909.127.6.797

Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer: New York, 2016.

Willadsen-Jensen EC, Ito TA. Ambiguity and the timecourse of racial perception. *Soc Cogn* 2006;**24**:580–606. https://doi.org/10.1521/soco.2006.24.5.580

Willadsen-Jensen EC, Ito TA. A foot in both worlds: Asian Americans' perceptions of Asian, White, and racially ambiguous faces. *Group Process Intergroup Relat* 2008;**11**:182–200.

Willadsen-Jensen E, Ito TA. The effect of context on responses to racially ambiguous faces: changes in perception and evaluation. *Soc Cogn Affect Neurosci* 2015;**10**:885–92. https://doi.org/10.1093/scan/nsu134

Zhang T, Zhou Y, Han S. Priority of racial and gender categorization of faces: a social task demand framework. *J Pers Soc Psychol* 2023;**124**:483–520. https://doi.org/10.1037/pspa0000318

Zhou Y, Gao T, Zhang T *et al*. Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nat Hum Behav* 2020;**4**:69–87. https://doi.org/10.1038/s41562-019-0743-y