Share the Wealth: Neurophysiological and Motivational Mechanisms Related to Racial Discrimination in Economic Decision Making

Hannah I. Volpert-Esmond[a]

Jessica R. Bray[b]

Meredith P. Levsen[c,d]

Bruce D. Bartholow[e]

[a]University of Texas at El Paso, Department of Psychology, 500 W. University Ave, El Paso TX 79912

[b]University of Kentucky, Department of Psychology, 503 Library Dr, Lexington KY 40508

[c]The New York Times, 620 8th Ave, New York NY 10018

[d]University of Missouri-Columbia, Department of Psychological Sciences, 200 S. 7th St, Columbia MO, 65211

[e]University of Iowa, Department of Psychological and Brain Sciences, 340 Iowa Ave, Iowa City IA, 52242

Accepted for publication at Journal of Experimental Social Psychology

Abstract

Social interactions are influenced by rapid judgements about interaction partners that are assumed to contribute to various behavioral biases. While often negligible in a given instance, such biases can accumulate to contribute to persistent inequities between social groups. Here, we used event-related potentials (ERPs) to determine the extent to which early attention to racial category information during simulated interpersonal interactions contributes to race bias in financial decisions. Undergraduate participants (N = 67; 36 women, 31 men; all White/Non-Hispanic) completed an economic decision-making task in which they decided how much money to invest in a series of male interaction partners (i.e., trustees) who varied in their apparent racial group memberships. Black male trustees received lower investments than White male trustees, replicating prior findings. Of greater interest, an ERP index of attention to trustees' faces predicted racial bias in investments, and was moderated by participants' internalized motivation to respond without prejudice (i.e., a difference score reflecting the extent to which participants' motivation reflected internal [e.g., personal egalitarian values] compared to external [e.g., concerns about social norms] reasons to respond without prejudice). Consistent with motivational models of prejudice control, greater early attention to a trustee's face led to *more*-biased lending among participants with lower internalized motivation but to *less*-biased lending among participants with higher internalized motivation. Findings demonstrate a crucial role for within-person variability in attention to race-related cues when interacting with others, along with between-person bias regulation motives, in determining whether attention to race will increase or decrease bias in financial lending.

Keywords*:* Racism, Psychophysiology, Behavioral Economics, Motivation, Trial-level modeling

Despite efforts to reduce racial discrimination and increase equity, a racial wealth gap persists in the United States, with Black families possessing $23 for every $100 held by White families (Kent & Rickets, 2024). Home equity is the largest contributor to wealth for most Americans (Schuetz, 2020) and is largely dependent on the ability to secure a mortgage loan. Although race-based discrimination in mortgage lending is prohibited, a recent analysis (Martinez & Kirchner, 2021) showed that Black Americans were 80% more likely to be denied a home mortgage than their White peers, even when differences in credit scores and other traditional risk indicators were accounted for. Ultimately, mortgage lending is one instance in which a lender makes social judgements about the applicant, wherein trust, cooperativeness, and prosocial decisions can be facilitated by a shared group identity (Tanis & Postmes, 2005). This often works against Black applicants, since Black Americans are significantly under-represented among loan officers (Frame et al., 2022). Addressing the persistent racial bias in lending requires better understanding of the mechanisms that contribute to the emergence of bias both within and between individuals. The present study tested the prediction that early attention to race, characterized with event-related brain potentials (ERPs), contributes to race bias in White participants' decisions to invest in White and Black male partners in an economic decision making game.

Race Bias in Economic Decision Making

Researchers have long been interested in how heuristic processes influence economic decisions (see Gigerenzer & Gaissmaier, 2011). Such studies often rely on decision-making games to investigate heuristic influences on resource allocation. For example, in the Trust Game (Berg et al., 1995; Camerer & Weigelt, 1988), the participant (i.e., investor) is given a financial stake prior to each trial and must decide how much to invest in an interaction partner (the

trustee). The investment is then quadrupled, and the trustee can either keep the entire amount or share half of it back with the investor, thereby doubling the investor's investment. Studies using the Trust Game consistently show a race bias, whereby White investors tend to invest less in Black relative to White trustees (Burns, 2006; 2012; Cañadas et al., 2015; Fershtman & Gneezy, 2001; Kubota et al., 2013; Stanley et al., 2011, 2012).

Variability in the magnitude of investment bias in the Trust Game has been linked to lenders' racial attitudes (Stanley et al., 2011; Tortosa et al., 2013). Such findings are consistent with a traditional between-person perspective on race bias, i.e., that some people are more biased than other people (e.g., Amodio et al., 2003; Greenwald et al., 1998). However, like other tendencies often assumed to be trait-like (Fleeson, 2004), the expression of bias also varies *within* persons (e.g., across time or situations)—and such variability is not well characterized in traditional approaches that distill responses across many trials into an average (see Vaughan & Birney, 2023). Moreover, although a large literature has characterized how varying features of target persons (e.g., racial prototypicality) affects racial category activation and/or bias (e.g., Johnson et al., 2015; Kawakami et al., 2017; Maddox et al., 2022), less is known regarding processes internal to the perceiver that can produce fluctuations in bias. Examining such factors is critical to better understanding circumstances in which bias is more or less likely to emerge.

Role of Attention to Race

Here, we investigated whether within-person variability in early attention to trustees' faces—and the racial category information they contain—contributes to biased investment decisions in the Trust Game. Salient features of visual stimuli elicit enhanced neuronal firing in visual cortex very early in stimulus processing (within ~100-200 ms), thereby commanding attention (Connor et al., 2004). Attended features are then selected for continued elaboration

(Tsotsos, 2005), biasing subsequent processing and influencing behavior (Zhaoping, 2005). The P2 component of the ERP, emerging ~150-200ms after stimulus onset at central midline scalp locations, is thought to reflect this early attention allocation process (e.g., Delplanque et al., 2004; Mangun, 1995; Schmitt et al., 2015). Social categories are known to be highly salient features that command attention during face processing (see Todorov, 2010), and a robust literature indicates that the P2 is highly sensitive to race (for reviews see Amodio & Cikara, 2021; Bartholow, 2023). For instance, Black faces consistently elicit larger P2 amplitudes than White faces (e.g., Amodio, 2010; Correll et al., 2006; Ito & Tomelleri, 2017; Ito & Urland, 2003, 2005; Kubota & Ito, 2007, 2017; Tortosa et al., 2013; Volpert-Esmond et al., 2017; Willadsen-Jensen & Ito, 2006, 2015). This pattern has been observed in White and Black perceivers in the U.S. (Volpert-Esmond & Bartholow, 2019, 2021; but see Dickter & Bartholow, 2007), in Chinese perceivers (Zhou et al., 2020), and in White European and North American perceivers living in China (Zhang et al., 2023).

Of importance here, early allocation of attention to social categories has implications for downstream behaviors. Across three experiments, Volpert-Esmond and Bartholow (2021) found that within-person variability in face-elicited P2 amplitude predicted the speed of social categorization, such that a larger P2 elicited by a given face—relative to the perceiver's average P2 amplitude—led to faster overt categorization of that face by race or gender. Consistent with continuous flow models of information processing (e.g., Coles et al., 1985; Requin et al., 1988), these findings support the idea that allocating greater attention to a face early in processing facilitates identification of goal-relevant features like race and gender, which also should facilitate activation of category-based knowledge (e.g., Rees et al., 2020) and trust (e.g., Cañadas et al., 2015; Foddy et al., 2009; Stanley et al., 2012; Tanis & Postmes, 2005).

The possibility that variability in early attention to faces contributes to variability in bias expression is suggested by prior work linking average P2 amplitude with average bias. Using data from a stereotype priming task (Payne, 2001), Amodio and Swencionis (2018) reported a positive association between the size of the "race effect" in P2 amplitude (i.e., larger to Black vs. White faces) and average levels of bias (also see Amodio, 2010; Correll et al., 2006), and that a manipulation intended to reduce attention to the face primes reduced both the average P2 race effect and average bias in behavior. A key limitation of these studies is their reliance on between-person averages, an approach that implicitly assumes responses are stable within persons and, hence, that only between-person variability is of interest. The present study extends this prior work by (i) examining associations between face-elicited P2 amplitude and behavior at the level of individual trials, shifting the focus from between-person averages to within-person variability, and (ii) applying this approach to understanding a behavior—investment decisions—that is more deliberative and complex than simple stimulus discrimination and that has more direct implications for understanding racial wealth disparities.

Role of Motivation to Respond Without Prejudice

Relative to general racial attitudes, individual differences in White people's motivations to respond without prejudice are important in the present context because they influence the quality of interracial interactions (see Butz & Plant, 2009; LaCosse & Plant, 2020). Specifically, White individuals with *internal* sources of motivation (e.g., important personal values to be unbiased) tend to focus on the quality of Black interaction partners' experiences and to show them respect, whereas those with *external* sources of motivation (e.g., wanting to avoiding social disapproval, concerns about being perceived as prejudiced) tend to focus on their own experiences and needs, thereby demonstrating less respect for their partner (LaCosse & Plant,

2020). More generally, relative to individuals high in external motivation and/or low in internal motivation, those with a combination of high internal motivation and low external motivation (i.e., more internalized motivation) display less race bias across a range of measures and situational contexts (e.g., Amodio et al., 2003, 2008a; Butz & Plant, 2009; Devine et al., 2002; Johns et al., 2008).

Of relevance here, the self-regulation of prejudice model (SRP; Monteith, 1993; Monteith et al., 2009) posits that attending to race cues has different consequences depending on perceivers' bias regulation motives. Among people whose motivation to respond without prejudice is highly internalized, focusing on race is theorized to activate the behavioral inhibition system, thereby facilitating control over the influence of racial stereotypes on behavior (see Monteith et al., 2002). Supporting this idea, Devine et al. (2002) reported that individuals high on internal and low on external motivation showed a reversal of the typical affective priming effect (i.e., slower reaction time) when classifying negative words following exposure to Black relative to White faces (also see Maddux et al., 2005). In contrast, among people for whom nonprejudiced responding is not highly internalized, the stereotypes activated by attending to race are not likely to be counteracted by inhibitory control and, thus, are more likely to bias responding. Here, we tested whether internalized motivation to respond without prejudice modulates the association between attention to race and racially-biased responding in real time, across a number of interactions between White investors and Black (vs. White) trustees.

Current Study

The aim of this study was to test whether within-person variability in investment bias can be explained, in part, by fluctuations in attention to race, represented here in trial-by-trial variability in P2 amplitude elicited by White and Black male trustee's faces. Our review of the

literature led to several predictions. First, we expected to replicate prior findings showing that, on average, White participants invest more in White relative to Black trustees (i.e., a main effect of trustee race on investment amount; see Stanley et al., 2011) and that these effects would be moderated by internalized motivation to respond without prejudice (i.e., an interaction between trustee race and motivation on investment amount). Specifically, we expected the average racial disparity in investments to decrease as a function of increased internalized motivation. Additionally, we expected to replicate prior findings that Black faces elicit larger P2 amplitudes than White faces (i.e., a main effect of trustee race on P2 amplitude; see Amodio & Cikara, 2021; Bartholow, 2023). Based on prior reports linking higher externalized motivation to an attention bias for Black relative to White faces (Bean et al., 2012; Richeson & Trawalter, 2008), we additionally expected the predicted effect of race on P2 amplitude to increase as a function of less internalized/more externalized motivation (i.e., an interaction between trustee race and motivation on P2 amplitude).

More pertinent to the aims of this study, and based on the idea that a greater focus on race in a given interaction is likely to induce stronger activation of racial stereotypes (Rees et al., 2020) and group-based differences in trust (e.g., Cañadas et al., 2015; Tanis & Postmes, 2005), we predicted that the magnitude of race bias in investments would vary as a function of within-person variability in the amplitude of the P2 elicited by trustees' faces (i.e., an interaction between trustee race and P2 amplitude on investment amounts). That is, we expected a larger P2 on a given trial, relative to a participant's average P2, to correspond with a larger investment when the trustee is White and a smaller investment when the trustee is Black. Finally, based on the theory that people for whom responding without prejudice is an important personal value use race as a cue to engage control over bias (e.g., Monteith et al., 2002, 2009), we expected this 2-

way interaction to be qualified by between-person differences in internalized motivation (i.e., a 3-way trustee race x P2 x motivation interaction on investment amounts). Specifically, we expected the model-predicted racial disparity in investments to *increase* as face-elicited P2 amplitude increased among individuals relatively lower in internalized motivation but predicted the opposite pattern—*decreased* racial disparity in investments as P2 amplitude increased—among individuals relatively higher in internalized motivation. We had no *a priori* predictions regarding the specific patterns of simple effects, as it was unclear whether internalized motivation would primarily affect investments in Black trustees, White trustees, or both.

Method

Participants

A convenience sample of 78 undergraduates (41 women, 37 men; ages 18-21, $M = 18.9$; $SD = 0.8$) enrolled in Introductory Psychology at [masked for review] participated in exchange for research participation credit. Participants also received up to $5 in cash depending on the outcome of two randomly selected trials from the Trust Game. The sample was primarily White/Non-Hispanic ($n = 67$), but also included Asian ($n = 1$), Black ($n = 7$), Hispanic ($n = 2$), and biracial ($n = 1$) individuals. Given our interest in White lenders' differential investments in Black versus White trustees, only data from White/Non-Hispanic participants were used for analysis. The targeted sample size ($n = 60$) was determined based on feasibility concerns and on comparable prior studies (e.g., Kubota et al., 2013; Stanley et al., 2011). No data were collected once data analysis commenced. A sensitivity power analysis using a summary-statistics-based approach (Murayama et al., 2022) determined that the smallest cross-level interaction that could be detected with 90% power with the final sample ($n = 67$) was $t = 3.3$ or Cohen's $d = 0.41$. All

procedures and materials used in this study were approved by the [masked for review] Internal Review Board (Project # 2001953).

Measures and Materials

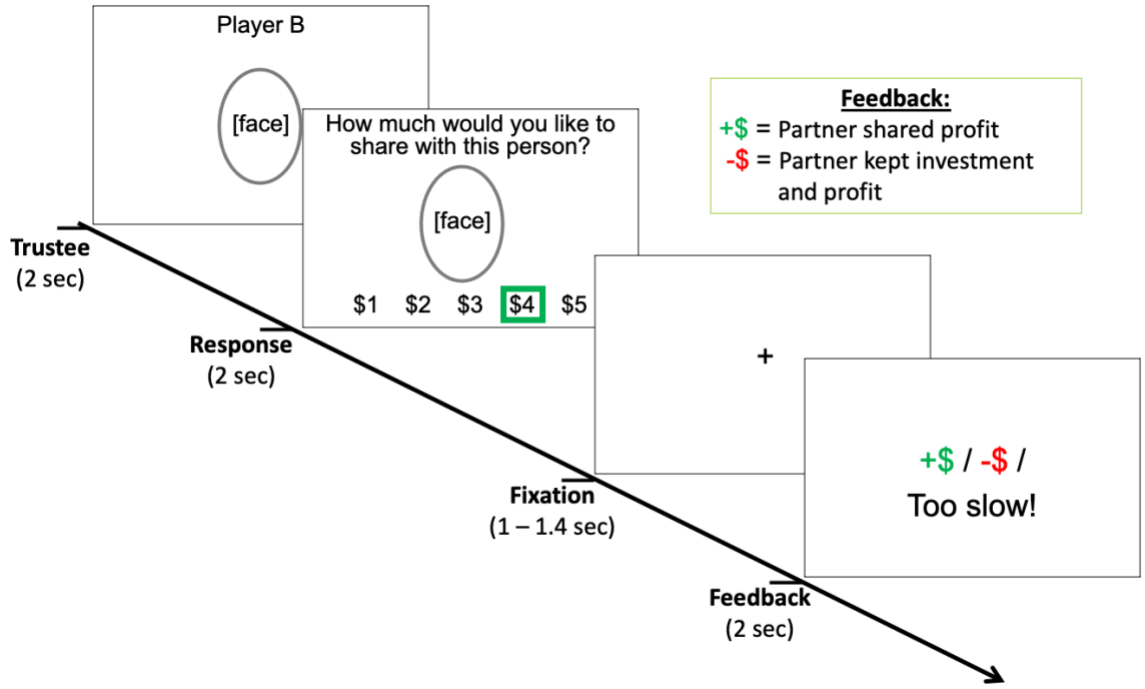
Trust Game

Participants played a series of "one-shot" trust games in which the trustee differed for each trial and was never repeated. At the beginning of each trial, participants received $5 to divide in whole number increments (from $1 to $5) between themselves and the trustee. They were told that the amount they shared would be quadrupled and that the trustee could either keep it all (initial investment, plus the profit) or share half of it back, thereby doubling the participant's investment. Participants were shown a picture of the trustee on each trial and were encouraged not to think too hard about their decisions but to rely more on their 'gut' (Kubota et al., 2013; Stanley et al., 2011, 2012).

Trial Timing. Trial format and timing are depicted in Figure 1. Each trial began with the words "Player B" at the top of the screen with a photo of the trustee (2 sec), followed by a screen reading, "How much would you like to share with this person?" Participants had 2 seconds to select an investment amount using a button box. Following the investment decision, a fixation cross was displayed for a variable duration (1, 1.2, or 1.4 sec), after which participants received visual feedback indicating whether the trustee had decided to share the profit with them ("+$" written in green) or to keep the investment and the profit ("-$" written in red). Investment choices that took longer than 2 sec elicited a "Too slow!" message instead. To reduce saccades, feedback text did not display the amounts won or lost (e.g., Tortosa et al., 2013).

Figure 1

Trial Procedure and Timing in the One-shot Trust Game



Note. 'Face' placeholder used in the figure to avoid including photographs of recognizable persons.

Participants completed five practice trials followed by 210 experimental trials divided into four blocks, separated by self-timed breaks. Of the 210 experimental trials, 80 trials displayed Black trustees, 80 displayed White trustees, and 50 displayed trustees of other races (e.g., Asian, Latino), included as fillers to obscure the primary purpose of the study (Kubota et al., 2013). Primary analyses used data from White and Black trustees; exploratory analyses of data from other-race trials are included in the Supplementary Material. Trials were presented in a random order and feedback outcomes occurred evenly within each racial group (e.g., 50% of the White trustee and Black trustee trials resulted in positive feedback).

Stimulus Pre-testing. Trustee photos were selected from 291 color pictures of men with neutral expressions (100 Black faces, 100 White faces, 91 "other race" faces) used in Stanley et

al. (2011).[1] A pre-testing sample ($N = 46$, recruited from the same pool as the main sample[2]) rated each of the faces on a scale ranging from 1 to 9, where a higher score represented more trustworthiness. These ratings were averaged within racial categories, and the subset of 210 face stimuli (80 Black, 80 White, 50 "other race" faces) used in the main study were selected to maintain the trust rating distributions within those categories (Ms = 4.46, 5.15, and 4.87 for Black, White, and other race faces, respectively). On average, Black faces were perceived as less trustworthy than White faces, $b = .69$, $t(45.0) = 5.1$, $p < .001$. To ensure that effects of trustee race were not confounded with group differences in perceived trustworthiness, analyses controlling for trustworthiness ratings were conducted and can be found in the Supplemental Material. (Patterns of effects were unchanged when controlling for trustworthiness ratings.)

Motivations to Respond Without Prejudice

Individual differences in motivations to respond without prejudice were assessed using the Internal and External Motivation to Respond Without Prejudice Scales (Plant & Devine, 1998; IMS and EMS, respectively). The IMS (5 items; $\alpha = .86$) measures the extent to which respondents attempt to respond without prejudice in their dealings with Black people because being egalitarian is personally important to them (e.g., "Being non-prejudiced toward Black people is important to my self-concept"). The EMS (5 items; $\alpha = .78$) measures the extent to which respondents attempt to respond without prejudice due to external pressures (e.g., "I try to hide any negative thoughts about Black people in order to avoid negative reactions from others"). Responses are made on a scale anchored at 1 (*Strongly disagree*) and 9 (*Strongly*

[1] Faces were drawn from the Karolinska Directed Emotional Faces, the Eberhardt Laboratory Face Database, the Color Facial Recognition Technology Database from the National Institute of Standards and Technology, and the NimStim Face Stimulus Set.

[2] Demographic characteristics of pilot participants were not assessed but given that they were recruited from the same pool as the main sample, their characteristics are assumed to be similar.

agree). IMS and EMS scores were not correlated, $r(65) = .08$, $p = .502$. Following prior studies (e.g., Johns et al., 2008), we created a relative measure of internalization by subtracting EMS scores from IMS scores and then standardizing across the sample. Higher scores on this difference score measure represent relatively more internalized and less externalized motivation to respond without prejudice (i.e., "internalized motivation"). One criticism of this kind of difference score approach is that it conflates individuals who score high on both scales with individuals who score low on both scales. However, IMS was negatively skewed in our sample, resulting in a relatively low prevalence of low-IMS/low-EMS scores, relative to high-IMS/high-EMS scores (see Supplementary Material). Because we had no *a priori* hypotheses regarding differences between high-IMS/high-EMS and low-IMS/low-EMS groups, we believed this to be an acceptable approach. Ancillary analyses using IMS and EMS scores as separate predictors are presented in the Supplementary Material and produced patterns of results very similar to those we report in the main text.

Electrophysiological Recording and Data Processing

The electroencephalogram (EEG) was recorded using 34 Ag/AgCl electrodes placed in standard locations and fixed in a stretch-lycra cap (ElectroCap, Inc., Eaton, OH). Scalp electrodes were referenced online to the right mastoid; an average mastoid reference was derived offline. Vertical and horizontal eye movements were recorded with additional electrodes placed 1 cm above and below the left eye and ~2 cm outside the outer canthus of each eye, respectively. Signals were amplified with a Synamps2 amplifier (Compumedics-Neuroscan, Charlotte, NC), sampled at 500 Hz and filtered online at 0.01-40 Hz. Ocular artifacts (i.e., blinks) were removed offline using a regression-based procedure (Semlitsch et al., 1986). Epochs of 1100 ms (including 100-ms pre-stimulus baseline), time-locked to face onset, were created for each trial.

Trials containing voltage deflections ± 75 microvolts (μV) within the epoch were rejected, as were no-response trials. The mean number of accepted trials per participant was 153.2 (max: 160 trials, min: 120 trials, *SD* = 7.9 trials). P2 amplitude was largest at centro-parietal midline electrodes (see Figure 2) and was quantified as the mean amplitude 135-185 ms following face onset at 15 centro-parietal scalp locations (C1, C2, C3, C4, Cz, CP1, CP2, CP3, CP4, CPz, P1, P2, P3, P4, and Pz).[3]

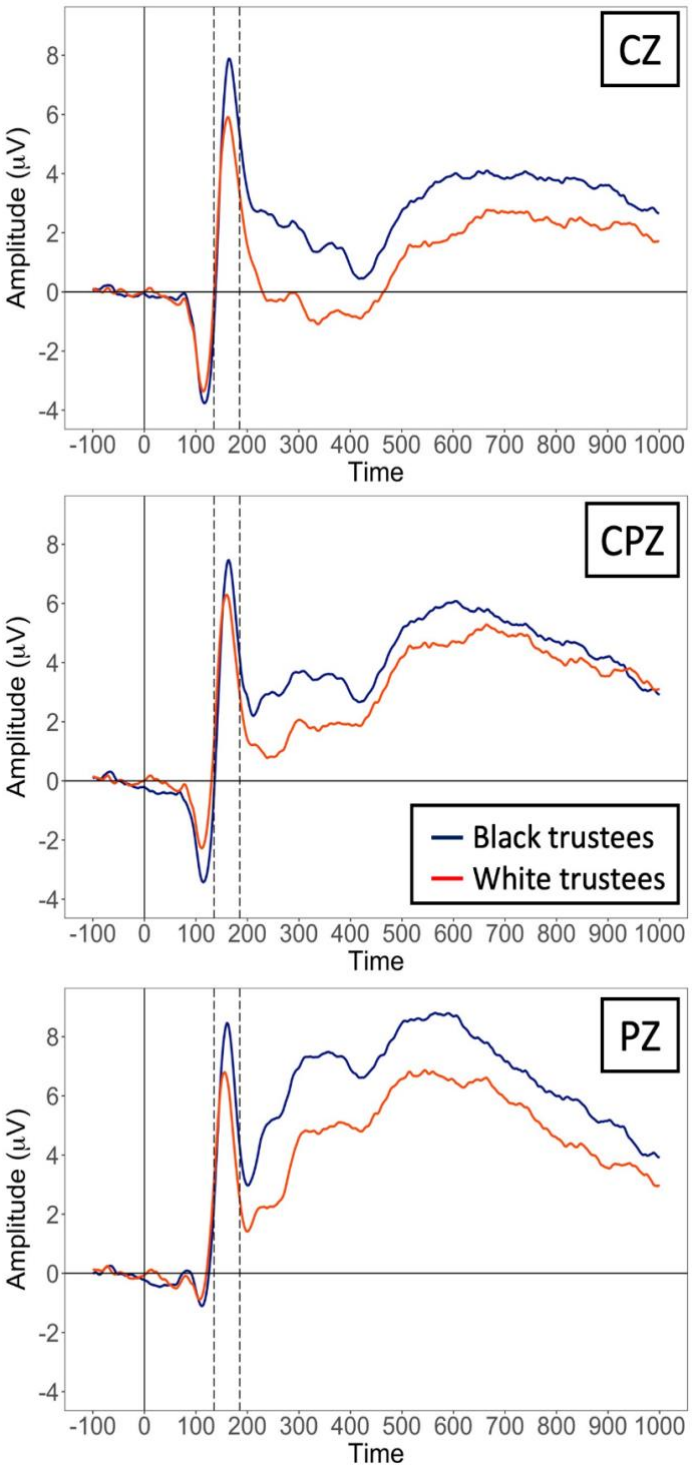
Procedure

Participants were told the study was part of an ongoing, large-scale project aimed at developing a large database of interaction outcomes during economic games used in psychology and economics, and that the individuals with whom they would be virtually "interacting" had

[3] In an exploratory manner, we also examined two other early-latency, face-elicited ERP components, the N1 and N2, which also are believed to reflect aspects of early modulation of attention (see Eason, 1981; Rugg et al., 1987). The N1 (85-135 ms), P2 (135-185 ms), and N2 (185-235 ms) components occur sequentially and can shed light on how attentional processes evolve over the course of face processing (see Ito & Urland, 2003). Because far less is known concerning how the amplitude of the N1 and N2 components relates to social categorization processes, results for these exploratory analyses are included in the Supplementary Material.

Figure 2

Grand Average ERP Waveforms Elicited by Trustees' Faces



Note. Dashed vertical lines mark the quantification window for the P2 (135-185 ms following face onset).

participated in the study previously. To bolster this cover story, they were asked to sign a second consent form and a photo release form (all agreed), after which a researcher took their photo. Next, EEG recording electrodes were applied and a researcher read the Trust Game instructions. Following the Trust Game, electrodes were removed, and participants were escorted to a private restroom to clean the electrode gel from their face and hair. Participants then were debriefed about the true nature of the experiment, thanked for their participation, and dismissed.

Data Analytic Approach

Hypotheses were tested using multilevel models (MLMs). MLMs confer several advantages over repeated-measures ANOVA for ERP and other intensively repeated-measures data (see Volpert-Esmond et al., 2021). For instance, MLMs allow for missing data, obviating the need for data imputation or listwise deletion of cases with missing observations on one or more variables. Additionally, MLMs allow specification of Level-1 predictors (e.g., ERP amplitude on a given trial) along with Level-2 predictors (e.g., traits of individual participants). This allows disaggregation of between-person and within-person variability within a predictor. Here, we applied the disaggregation method proposed by Curran and Bauer (2011) to separate trial-level P2 amplitude into between-person (average P2 amplitude, derived from all Black and White trustees' faces) and within-person (P2 amplitude to each face, centered around that participant's average) predictors, the latter being of primary interest for our hypotheses.

Since all data is modeled at the trial level, it is difficult to account for "bias" within each individual trial because bias requires a comparison across trials (e.g., comparing Black-trustee trials and White-trustee trials). Here, we operationalize racial bias as differences in model-estimated levels of the dependent variable (e.g., investment amount) by trustee race at given levels of other predictors (e.g., within-person P2 amplitude). For example, race bias in

investments is operationalized as the investment amount on a trial when the trustee is White and his face elicits a P2 = *x*, relative to the investment amount on a trial when the trustee is Black and his face elicits a similar P2 (i.e., *x*). Another way of conceptualizing this is simply as the model-estimated effect of trustee race on the dependent variable (and how the effect of trustee race is qualified by higher order interactions). In this way, the MLM approach permits trial-level inferences from data acquired across trials.

All models were fitted to trial-level data using the R package 'lme4' (Bates et al., 2015). The covariance structure for random effects was unstructured in all models. The most complex random effects structure supported by the data was used in each model, starting with a maximal model and systematically removing highly correlated random slopes (Matuschek et al., 2017). Satterthwaite approximations were used to estimate degrees of freedom (*df*) and to obtain two-tailed *p*-values; in situations where *df* >200, results are reported as *z* statistics. All reported means are estimated marginal means (Lenth, 2022). Internalized motivation was standardized (centered and scaled) across the sample to assist with interpretation. Effect size within 2-level multilevel models is calculated using a summary-statistics-based approach, where Cohen's *d* is used as an effect size for Level-1 effects and Pearson's *r* is used as a measure of effect size for Level-2 effects and cross-level interactions (Murayama et al., 2022).

All data and code used for analyses can be viewed at [anonymized link: https://osf.io/g7wjb/?view_only=c0f586fc603d4449806c6229da9d79b7]. Additional data, including raw EEG files, are available upon request. All studies, experimental manipulations, and exclusions are reported. Additional self-report questionnaires administered but not examined in the main text are reported in the Supplementary Material.

Results

Racial Bias in Investments

The average investment in trustees across participants was $2.78 (*SD* = $1.20). The first model examined how investment amounts varied as a function of trustee race and internalized motivation (and their interaction; see Table 1).[4] Consistent with prior research (e.g., Stanley et al., 2011), participants invested more in White male trustees (*M* = $2.88, *SE* = $0.09) than in Black male trustees (*M* = $2.69, *SE* = $0.07). There was no main effect of internalized motivation. The interaction of trustee race and internalized motivation also was not significant, but the pattern of means was as hypothesized, i.e., more equivalent investments in White and Black trustees among participants whose motivation to respond without prejudice was more internalized.

Table 1

Fixed Effects from the Multilevel Model Predicting Investment Amounts as a Function of Trustee Race and Internalized Motivation

Predictor	*b*	*df*	*t*	*p*	Effect size
Intercept	2.69	82.8	29.53	<.001	
Trustee race	0.19	129.9	2.50	.014	0.31
Motivation	0.00	65.0	0.05	.956	0.01
Trustee race x Motivation	-0.12	65.0	-1.90	.062	0.23

Attention to Trustees' Faces as a Function of Race and Motivation

[4] The data structure was at the trial level (i.e., predicting investment amounts on each trial). Trustee race was included as a random slope within participant and a random intercept was included for face stimulus. Wilkinson notation: `InvestmentAmount ~ TrusteeRace*Motivation + (TrusteeRace|Participant) + (1|FaceStimulus)`

The second model examined predictors of P2 amplitude on each trial (see Table 2).[5] As in prior research (see Amodio & Cikara, 2021), on average, Black male faces elicited larger P2s (*M* = 5.56 μV, *SE* = 0.49) than White male faces (*M* = 4.38 μV, *SE* = 0.47). Neither the effect of internalized motivation to respond without prejudice nor its interaction with trustee race was significant; the patterns represented in the interaction were such that more internalized motivation corresponded with a larger racial disparity in P2 amplitude, which opposes the hypothesized pattern (based on Bean et al., 2012; Richeson & Trawalter, 2008).

Table 2

Fixed Effects from the Multilevel Model Predicting P2 Amplitude as a Function of Trustee Race and Internalized Motivation

Predictor	*b*	*df*	*t*	*p*	Effect size
Intercept	5.56	98.1	11.44	<.001	
Trustee race	-1.18	196.0	-3.80	<.001	-0.47
Motivation	-0.65	65.0	-1.52	.133	0.19
Trustee race x Motivation	-0.35	65.1	-1.75	.085	0.21

Investment Amount as a Function of Attention, Race, and Motivation

The third model tested our central hypothesis, i.e., that within-person variability in attention to trustees' faces (P2 amplitude) would predict the amount invested on each trial, and that this predicted association would vary according to trustee race and differences in internalized motivation (see Table 3).[6] Unsurprisingly, between-person variability in (average)

[5]The data structure was at the trial level (i.e., predicting P2 amplitude on each trial). Trustee race, internalized motivation, and their interaction were all included as fixed predictors. A random slope for trustee race was included within participant and random intercepts were included for electrode and face stimulus. Wilkinson notation: `P2 ~ TrusteeRace*Motivation + (TrusteeRace|Participant) + (1|Electrode) + (1|FaceStimulus)`

[6]The data structure was at the trial level (i.e., predicting investment amounts on each trial). Participant-centered P2 amplitude (i.e., within-participant variability in P2 amplitude), trustee race, internalized motivation, and their

P2 amplitude was unrelated to investments (see Volpert-Esmond & Bartholow, 2021). As predicted, however, the model produced a significant P2 (within-person) x Trustee race interaction: as face-elicited P2 amplitude on a given trial increased, so too did the model-predicted discrepancy in the amount invested in White versus Black male trustees. Post-hoc tests showed that whereas investments in White trustees increased if their faces elicited larger P2 amplitudes (relative to a participant's average P2), b = .0014 (95% CI: .0007; .0020), investments in Black trustees were unrelated to the size of the P2s their faces elicited, b = .0004 (95% CI: -.0003; .0011).

Table 3

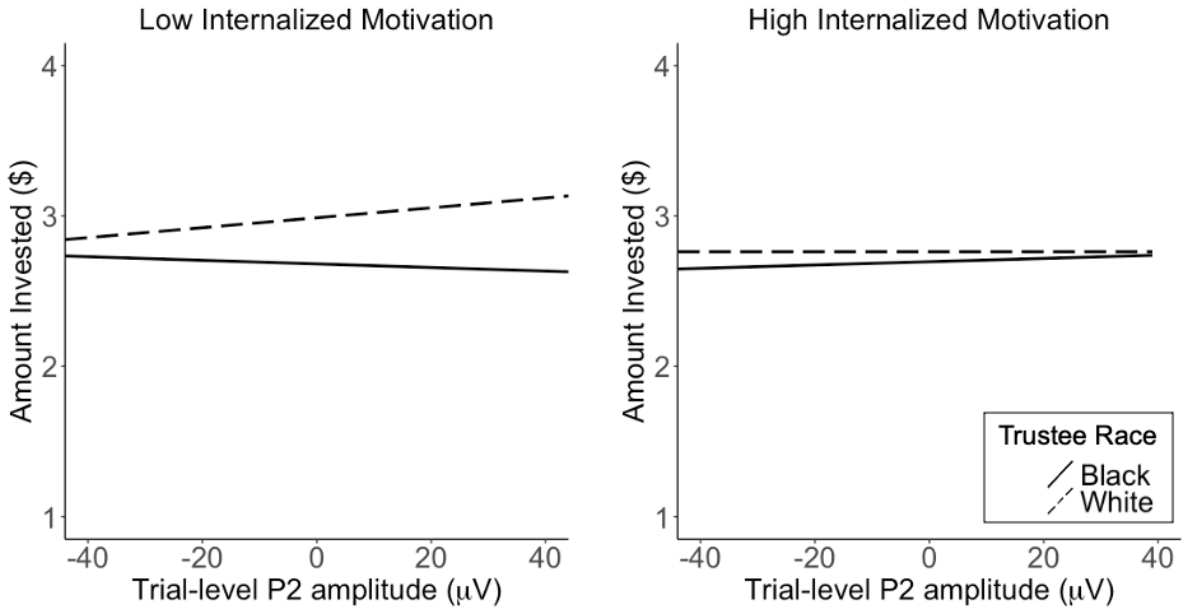
Fixed Effects from the Multilevel Model Predicting Investment Amounts as a Function of P2 Amplitude, Trustee Race, and Internalized Motivation

Predictor	*b*	*df*	*t*	*p*	Effect size
Intercept	2.686	88.7	28.91	<.001	
P2 (between)	0.000	>200	0.05	.963	0.01
P2 (within)	0.000	>200	1.08	.279	0.13
Trustee race	0.189	146.2	2.41	.017	0.30
Motivation	0.005	65.1	0.06	.956	0.01
P2 (within) x Trustee race	0.001	>200	2.02	.044	0.25
P2 (within) x Motivation	0.001	>200	3.60	<.001	0.41
Trustee race x Motivation	-0.117	64.8	-1.90	.062	0.23
P2 (within) x Trustee race x Motivation	-0.003	>200	-6.12	<.001	0.60

interactions were all included as fixed predictors, as well as mean P2 amplitude (i.e., between-participant variability) in P2 amplitude. A random slope for trustee race was included within participant and random intercepts were included for electrode and face stimulus. Wilkinson notation: `InvestmentAmount ~ P2.between + P2.within*TrusteeRace*Motivation + (TrusteeRace|Participant) + (1|Electrode) + (1|FaceStimulus)`

Figure 3

Investment Amounts as a Function of Trial-level P2 Amplitude, Trustee Race, and Internalized Motivation



Note. "Low" and "High" Internalized Motivation represent values -1 and +1 *SD* from the mean of that variable, respectively. Trial-level P2 amplitude was centered around each individual's average P2 amplitude, providing an index of within-person variability in P2 amplitude from trial to trial.

This two-way interaction was further qualified by the predicted P2 (within) x Trustee Race x Motivation interaction (see Figure 3). To unpack the interaction, we separately examined the form of the Trustee race x P2 amplitude interaction at low (-1 *SD*) and high (+1 *SD*) levels of internalized motivation. Among individuals relatively low on internalized motivation (i.e., left panel in Figure 3), greater allocation of attention to the trustee's face on a particular trial, relative to a participant's own average P2 amplitude, was positively related to the amount they invested in White male trustees, $b = .0031$ (95% CI: 0.0022; 0.0040), but was unrelated to the amount they invested in Black male trustees, $b = -.0009$ (95% CI: -0.0020; 0.0001). These slope differences correspond to an increase in the racial disparity in investments as attention to race

increased (see Table 4). In contrast, for individuals relatively high on internalized motivation (i.e., right panel in Figure 3), greater allocation of attention to the trustee's face increased the amount participants invested in Black male trustees, $b = 0.0016$ (95% CI: 0.0006; 0.0026) but had no effect on the amount invested in White male trustees, $b = -0.0004$ (95% CI: -0.0013; 0.0006). These slope differences correspond to a *decrease* in the racial disparity in investments as attention to race increased (see Table 4).

Table 4

Post-hoc, Model-estimated Mean Investments at Low and High Internalized Motivation to Respond Without Prejudice, as a Function of Small and Large Relative P2 Amplitudes

	Low Internalized Motivation	High Internalized Motivation
Small P2 amplitude (M = -10.69 μV)	Black trustees: \$2.69 (SE = 0.13) White trustees: \$2.96 (SE = 0.11) Contrast: b = -0.26, SE = 0.1, p = .008	Black trustees: \$2.67 (SE = 0.13) White trustees: \$2.77 (SE = 0.11) Contrast: b = -0.09, SE = 0.1, p = .351
Large P2 amplitude (M = 10.69 μV)	Black trustees: \$2.67 (SE = 0.13) White trustees: \$3.02 (SE = 0.11) Contrast: b = -0.35, SE = 0.1, p = .001	Black trustees: \$2.71 (SE = 0.13) White trustees: \$2.76 (SE = 0.11) Contrast: b = -0.05, SE = 0.1, p = .613

Note. "Low" and "High" values for Internalized Motivation, and "Small" and "Large" relative P2 amplitudes, represent -1 and +1 *SD* values, respectively.

Discussion

The present study provides the first evidence that naturally occurring, within-person fluctuations in early allocation of attention to faces, shown in prior work to predict the efficiency of overt race and gender categorization (Volpert-Esmond & Bartholow, 2021), contributes meaningfully to the behavioral expression of race bias in financial decisions. Allocation of attention to a trustee's face within 200 ms, as indexed by the amplitude of the face-elicited P2, explained significant variance in the amount of money invested in that trustee. More specifically,

greater attention to a given White male trustee's face, relative to the individual participant's average level of attention to faces, corresponded with increased investment in that trustee. This relationship was not observed for Black male trustees, which is consistent with a pattern of ingroup favoritism in this sample of White undergraduates (see Brewer, 2017).

However, this general pattern crucially depended upon participants' internalized motivation for responding without prejudice, in a manner consistent with theory. Monteith's (1993) self-regulation of prejudice model holds that, for individuals motivated by personal values to respond without prejudice, the perception of race serves as a cue for the possibility of bias, thereby engaging inhibitory control over behavior (Monteith et al., 2002). The present findings are in-line with this reasoning. Among individuals for whom the motivation to respond without prejudice was more internalized, a larger P2 elicited by a given Black male trustee's face (relative to the participant's average P2 response to all Black and White trustees' faces) corresponded with increased investment in that trustee, a pattern not observed for White male trustees. As illustrated in Figure 3, this response profile corresponded with a pattern in which increased attention to trustees' faces was associated with decreased racial bias in investments.

In contrast, among participants for whom the motivation to respond without prejudice was less internalized, an opposing pattern emerged in which a larger relative P2 amplitude elicited by a given White male trustee's face was related to a larger investment in that trustee, whereas this association was absent for Black male trustees—effectively *increasing* racial bias in investments as attention to the trustee's face increased (see Figure 3; Table 4). In other words, the general pattern of increased attention to race facilitating ingroup favoritism was evident among participants whose bias control motives were less internalized but was absent among participants whose motives were more internalized, for whom attending to race produced a

modest (but significant) tendency to favor outgroup trustees. This finding extends prior work (Amodio, 2010; Amodio & Swencionis, 2018) by (i) linking the face-elicited P2 to patterns of ingroup and outgroup favoritism that differ according to levels of internalized motivation, (ii) demonstrating a link between face-elicited P2 and race bias in a more deliberative behavior with implications for understanding the racial wealth gap, and, critically, (iii) demonstrating this phenomenon at the level of individual trials (i.e., individual interactions).

Implications of Within-person Variability in Categorization and Bias

Research on race bias has been dominated by perspectives that focus on (1) individuals' racial attitudes, either emphasizing sameness (i.e., all people in a given culture are consistently biased; see Dasgupta, 2009; Pauker et al., 2022; Weisbuch et al., 2009) or between-person differences (e.g., Fazio et al., 1995; Greenwald et al., 1998); and (2) individuals' ability to exert control over activated stereotypes (e.g., Gonsalkorale et al., 2009; Ito et al., 2015; Sherman et al., 2008). Most research using ERP and other neurophysiological measures to understand bias focus similarly on average responses at the level of the individual (e.g., Amodio, 2010; Amodio & Swencionis, 2018; Ito et al., 2004; Phelps et al., 2000; Stanley et al., 2012). The present findings underscore the importance of *within*-person differences in the extent to which processes that summon bias—or efforts to regulate it—are initiated. Such naturally occurring fluctuations in neural and behavioral responses across trials generally are treated as nuisance or error variance in both experimental and individual-differences research that relies on averaging across trials (Luck, 2014; Volpert-Esmond et al., 2018). We and others have shown in prior research that, when analyzed properly (e.g., with well-specified multilevel models) using carefully acquired and clean data, trial-level differences in ERP and behavioral responses represent meaningful variability in signals of interest and are not simply noise (see Kristjansson et al., 2007; Page-

Gould, 2017; Ratcliff et al., 2009; Volpert-Esmond et al., 2018, 2021; Von Gunten et al., 2018). Arguably, relative to standard between-person approaches, a focus on within-person, trial-by-trial variability may better represent the ways in which mechanisms of bias regulation operate in the natural environment, where numerous internal and external factors impinge upon attention to others' social category cues. Thus, this approach to examining laboratory responses could have improved ecological validity for understanding bias in the real world (see Andersen et al., 2023; Salmon & Hehman, 2022).

That assignment of others to social categories does not occur in a discrete, all-or-nothing fashion is by now well understood (see Freeman & Ambady, 2011; Johnson & Freeman, 2010). Moreover, that variability in social categorization processes has implications for bias is also unremarkable, having been demonstrated many times (see Freeman & Johnson, 2016; Johnson et al., 2015; Maddox et al., 2022). However, the existing literature on these topics primarily comprises demonstrations that categorization and expressions of bias can vary according to properties of the targets of perception (e.g., racial phenotypicality vs. ambiguity; see Maddox et al., 2022) or the contexts in which they are perceived (see Freeman et al., 2020; Krosch & Amodio, 2014). In contrast, the present findings emphasize that some process(es) internal to the perceiver that varies over the course of an experiment yields differing degrees of attention to racial category information, expressed in the magnitude of a very rapidly unfolding neurophysiological response to faces (also see Volpert-Esmond & Bartholow, 2017, 2019, 2021). That the magnitude of this early-latency, categorization-related response to a given face shapes behavior toward that target person *is* remarkable and has implications for theories related to how person construal leads to bias (e.g., see Kawakami et al., 2017; Petsko & Bodenhausen, 2020).

The basic premise that earlier access to relevant information facilitates subsequent classification decisions was established long ago by studies supporting continuous flow conceptualizations of information processing (e.g., Coles et al., 1985; Eriksen & Schultz, 1979; Eriksen et al., 1985; see also Erb et al., 2021). More recently, this basic premise was extended to social categorization as feature of the dynamic interactive model of person construal (Freeman & Ambady, 2011). Yet, in a strict sense, such models are limited to situations in which a target must be identified (i.e., "Who or what is that?"), often in the presence of irrelevant or distracting information. The present findings extend such models by suggesting that early access to target-related information affects more complex interpersonal decisions (e.g., "What will I do?") that have implications for the perceiver's own outcomes under conditions of uncertainty. In doing so, these findings could suggest ways to incorporate models of the role of attention in risky decision scenarios (e.g., Brandstätter & Körner, 2014; Johnson & Busemeyer, 2016; Zilker & Pachur, 2023) with models of person construal and its consequences.

Conclusions and Future Directions

Despite efforts to reduce racial discrimination and increase equity, household wealth held by Black families in the U.S. remains a fraction of that held by White families, a disparity that has changed very little in recent decades (Kent & Ricketts, 2024; Kent et al., 2020; Kraus et al., 2017). By elucidating the role of within-person variability in early attention to race cues and its implications for behavior, the current research contributes to understanding potential mechanisms by which lending decisions—critically important to the accumulation of wealth (Schuetz, 2020)—might be biased on some occasions more than others.

Future research in this area should endeavor to identify the psychological mechanism(s) linking early attention to racial categories with biased investment decisions. Several lines of prior

research point to threat perception as one candidate mechanism. In the context of economic decisions, and especially under conditions of scarcity (see Krosch et al., 2017), outgroup members are threatening to the extent that they have control over resources (see Esses et al., 1998; LeVine & Campbell, 1972). The trust game puts control over resources in the hands of trustees who determine whether to keep or share participants' financial investments, potentially elevating perception of threat in outgroup trustees. Thus, for a White perceiver, an elevated P2 elicited by a Black trustee's face could signal concerns about outgroup control over finite resources, thereby leading to lower investment in that trustee. Moreover, Krosch et al. (2017) demonstrated that, when resources are scarce, individuals low in internalized motivation to respond without prejudice allocate fewer resources to Black recipients, whereas the opposite pattern emerges among individuals high in internalized motivation. Together, these findings suggest that variability in the extent to which concerns about an outgroup trustee's control over monetary resources could account for the relationship between P2 amplitude and investment decisions, as well as the different form of this association for perceivers lower versus higher in internalized motivation.

Importantly, given that gender and race intersect in important ways to predict bias and discrimination (Carastathis, 2014), the generalizability of the present findings is limited by our use of only male trustees. Especially in financial interactions based on trust, feelings of threat elicited by outgroup men relative to outgroup women (Navarette et al., 2010)—as well as specific stereotypes associated with Black men and Black women (Melson-Silimon et al. 2024)—may be particularly important to consider. Additionally, some limited research highlights differential patterns of attention to faces when perceivers' explicit goals are to categorize race versus gender (e.g., Ito & Urland 2003), and that race and gender interact in predicting the

allocation of attention (Ito & Urland, 2005; Volpert-Esmond & Bartholow, 2019). To address this limitation, future research should examine how trustee gender may play a role in racially biased investment decisions.

In conclusion, the present findings extend prior research on the mechanisms of the racial disparity in financial investments in White compared to Black male trustees during economic games (e.g., Kubota et al., 2013; Stanley et al., 2011; Tortosa et al., 2013). More broadly, these findings support and extend models of intergroup bias by highlighting the interplay between within-person variability in early attention to race and between-person differences in chronic prejudice control-related motivations. Future research should endeavor to determine the generalizability of these patterns for other forms of bias and to identify the internal and external factors that contribute to fluctuations in attention to social category information represented in others' faces.

Open Practices

The experiment in this article requests consideration for the Open Data badge for transparent practices. All data and code for analyses are available at [anonymized link: https://osf.io/g7wjb/?view_only=c0f586fc603d4449806c6229da9d79b7].

References

Amodio, D. M. (2010). Coordinated Roles of Motivation and Perception in the Regulation of Intergroup Responses: Frontal Cortical Asymmetry Effects on the P2 Event-related Potential and Behavior. *Journal of Cognitive Neuroscience*, *22*(11), 2609–2617. https://doi.org/10.1162/jocn.2009.21395

Amodio, D. M., Bartholow, B. D., & Ito, T. A. (2014). Tracking the dynamics of the social brain: ERP approaches for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *9*(3), 385–393. https://doi.org/10.1093/scan/nst177

Amodio, D. M., & Cikara, M. (2021). The Social Neuroscience of Prejudice. *Annual Review of Psychology*, *72*(1). https://doi.org/10.1146/annurev-psych-010419-050928

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, *18*(6), 524–530. https://doi.org/10.1111/j.1467-9280.2007.01933.x

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008a). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, *94*(1), 60–74. https://doi.org/10.1037/0022-3514.94.1.60

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84*(4), 738–753. https://doi.org/10.1037/0022-3514.84.4.738

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2008b). Neurocognitive components of the behavioral inhibition and activation systems: Implications for theories of self-

regulation. *Psychophysiology, 45*, 11–19. https://doi.org/10.1111/j.1469-8986.2007.00609.x

Amodio, D. M., & Swencionis, J. K. (2018). Proactive control of implicit bias: A theoretical model and implications for behavior change. *Journal of Personality and Social Psychology, 115*(2), 255–275. https://doi.org/10.1037/pspi0000128

Andersen, J. P., Di Nota, P. M., Boychuk, E. C., Schimmack, U., & Collins, P. I. (2023). Racial bias and lethal force errors among Canadian police officers. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement, 55*(2), 130–141. https://doi.org/10.1037/cbs0000296

Bartholow, B.D. (2023). Default categorization of outgroup faces and the other race effect: Commentary on the special issue. *British Journal of Psychology,* 114 (Suppl. 1), 10-13. https://doi.org/10.1111/bjop.12648

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bean, M. G., Slaten, D. G., Horton, W. S., Murphy, M. C., Todd, A. R., & Richeson, J. A. (2012). Prejudice concerns and race-based attentional bias: New evidence from eye-tracking. *Social Psychological and Personality Science*, *3*(6), 722-729. https://doi.org/10.1177/1948550612436983

Berg, J., Dickhaut, D., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142. https://doi.org/10.1006/game.1995.1027

Brandstätter, E., & Körner, C. (2014). Attention in risky choice. *Acta psychologica*, *152*, 166–176. https://doi.org/10.1016/j.actpsy.2014.08.008

Brewer, M. B. (2017). Intergroup discrimination: Ingroup love or outgroup hate? In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 90–110). New York: Cambridge University Press. https://doi.org/10.1017/9781316161579.005

Burns, J. (2006). Racial stereotypes, stigma and trust in post-apartheid South Africa. *Economic Modelling*, *23*(5), 805–821. https://doi.org/10.1016/j.econmod.2005.10.008

Burns, J. (2012). Race, diversity and pro-social behavior in a segmented society. *Journal of Economic Behavior & Organization, 81*, 366-378. https://doi.org/10.1016/j.jebo.2011.11.006

Butz, D. A., & Plant, E. A. (2009). Prejudice control and interracial relations: The role of motivation to respond without prejudice. *Journal of Personality*, *77*(5), 1311-1342. https://doi.org/10.1111/j.1467-6494.2009.00583.x

Camerer, C., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica*, *56*(1), 1-36. https://doi.org/10.2307/1911840

Cañadas, E., Rodríguez-Bailón, R., & Lupiáñez, J. (2015). The effect of social categorization on trust decisions in a trust game paradigm. *Frontiers in Psychology*, *6*, 1568. https://doi.org/10.3389/fpsyg.2015.01568

Carastathis, A. (2014). The concept of intersectionality in feminist theory. *Philosophy Compass, 9*(5), 304-314. https://doi.org/10.1111/phc3.12129

Coles, M. G., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology. Human Perception and Performance*, *11*(5), 529–553. https://doi.org/10.1037//0096-1523.11.5.529

Connor, C.E., Egeth, H.E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology, 14*, R850-R852. https://doi.org/10.1016/j.cub.2004.09.041

Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, *42*(1), 120–128. https://doi.org/10.1016/j.jesp.2005.02.006

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Dasgupta, N. (2009). Color lines in the mind: Implicit prejudice, discrimination, and the potential for change. In A. Grant-Thomas & G. Orfield (Eds.), *Color lines: Exploring the frontiers of America's multiracial present and future* (pp. 97–117). Philadelphia, PA: Temple University Press.

Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., & Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neuroscience Letters*, *356*(1), 1–4. https://doi.org/10.1016/j.neulet.2003.10.014

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*(5), 835-848. https://doi.org/10.1037/0022-3514.82.5.835

Dickter, C. L., & Bartholow, B. D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Social Cognitive and Affective Neuroscience, 2*(3), 189-198. https://doi.org/10.1093/scan/nsm012

Eason, R. G. (1981). Visual evoked potential correlates of early neural filtering during selective attention. *Bulletin of the Psychonomic Society, 18*(4), 203–206. https://doi.org/10.3758/BF03333604

Erb, C. D., Smith, K. A., & Moher, J. (2021). Tracking continuities in the flanker task: From continuous flow to movement trajectories. *Attention, Perception & Psychophysics*, *83*(2), 731–747. https://doi.org/10.3758/s13414-020-02154-4

Eriksen, C. W., Coles, M. G., Morris, L. R., & O'Hara, W. P. (1985). An electromyographic examination of response competition. *Bulletin of the Psychonomic Society, 23*(3), 165–168. https://doi.org/10.3758/BF03329816

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*(4), 249–263. https://doi.org/10.3758/BF03198804

Esses, V. M., Jackson, L. M., & Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues, 54*, 699 – 724.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013–1027. https://doi.org/10.1037/0022-3514.69.6.1013

Fershtman, C., & Gneezy, U. (2001). Discrimination in a Segmented Society: An Experimental Approach. *The Quarterly Journal of Economics*, *116*(1), 351–377. https://doi.org/10.1162/003355301556338

Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science, 13*(2), 83–87. https://doi.org/10.1111/j.0963-7214.2004.00280.x

Frame, W. S., Huang, R., Mayer, E. J., & Sunderam, A. (2022, June 21). The impact of minority representation at mortgage lenders. Working paper, Federal Reserve Bank of Dallas (2207). https://doi.org/10.24149/wp2207

Foddy, M., Platow, M. J., & Yamagishi, T. (2009). Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science*, *20*(4), 419-422.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Freeman, J. B., & Johnson, K. L. (2016). More than meets the eye: Split-second social perception. *Trends in Cognitive Sciences*, *20*(5), 362–374. https://doi.org/10.1016/j.tics.2016.03.003

Freeman, J.B., Stolier, R.M., Brooks, J.A. (2020). Dynamic interactive theory as a domain-general account of social perception. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 61, pp. 237-287). Academic Press.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*(1), 451-482. https://doi.org/10.1016/bs.aesp.2019.09.005

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464-1480. https://doi.org/10.1037//0022-3514.74.6.1464

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology, 45*(2), 410–414. https://doi.org/10.1016/j.jesp.2008.11.004

Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, *13*(12), 524–531. https://doi.org/10.1016/j.tics.2009.10.002

Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, *108*(2), 187-218. https://doi.org/10.1037/a0038557

Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the timecourse of social perception: the effects of racial cues on event-related brain potentials. *Personality & Social Psychology Bulletin*, *30*(10), 1267–1280. https://doi.org/10.1177/0146167204264335

Ito, T. A., & Tomelleri, S. (2017). Seeing is not stereotyping: the functional independence of categorization and stereotype activation. *Social Cognitive and Affective Neuroscience*, *12*(5), 758–764. https://doi.org/10.1093/scan/nsx009

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*(4), 616–626. https://doi.org/10.1037/0022-3514.85.4.616

Ito, T.A., & Urland, G. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 21-36. https://doi.org/10.3758/CABN.5.1.21

Johns, M., Cullum, J., Smith, T., & Freng, S. (2008). Internal motivation to respond without prejudice and automatic egalitarian goal activation. *Journal of Experimental Social Psychology*, *44*(6), 1514-1519. https://doi.org/10.1016/j.jesp.2008.07.003

Johnson, J. G., & Busemeyer, J. R. (2016). A computational model of the attention process in risky choice. *Decision, 3*(4), 254–280. https://doi.org/10.1037/dec0000050

Johnson, K.L., & Freeman, J.B. (2010). A "new look" at person construal: seeing beyond dominance and discreteness. In E. Balcetis & D. Lassiter (Eds.), *The social psychology of visual perception* (pp. 253-272). New York: Psychology Press.

Johnson, K. L., Lick, D. J., & Carpinella, C. M. (2015). Emergent research in social vision: An integrated approach to the determinants and consequences of social categorization. *Social and Personality Psychology Compass, 9*(1), 15–30. https://doi.org/10.1111/spc3.12147

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. In J. M. Olson (Ed.), *Advances in experimental social psychology* (pp. 1–80). Elsevier Academic Press. https://doi.org/10.1016/bs.aesp.2016.10.001

Kent, A.H., & Ricketts, L.R. (2024, August). The state of U.S. wealth inequality. Federal Reserve Bank of St. Louis. https://www.stlouisfed.org/institute-for-economic-equity/the-state-of-us-wealth-inequality

Kent, A. H., Ricketts, L. R., Emmons, W. R., & Boshara, R (2020, November). The real state of family wealth: Will COVID-19 worsen racial, educational and generational Gaps in the U.S.? Federal Research Bank of St. Louis. https://www.stlouisfed.org/on-the-economy/2020/november/real-state-family-wealth-covid19-racial-gaps

Kraus, M. W., Rucker, J. M., & Richeson, J. A. (2017). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences, 114*(39), 10324-10331. https://doi.org/10.1073/pnas.1707719114

Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: an introduction to growth curve modeling. *Psychophysiology*, *44*(5), 728–736. https://doi.org/10.1111/j.1469-8986.2007.00544.x

Krosch, A. R., & Amodio, D. M. (2014). Economic scarcity alters the perception of race. *Proceedings of the National Academy of Sciences*, *111*(25), 9079-9084. https://doi.org/10.1073/pnas.1404448111

Krosch, A. R., Tyler, T. R., & Amodio, D. M. (2017). Race and recession: Effects of economic scarcity on racial discrimination. *Journal of Personality and Social Psychology*, *113*(6), 892-909. https://doi.org/10.1037/pspi0000112

Kubota, J. T., & Ito, T. A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*(5), 738–752. https://doi.org/10.1016/j.jesp.2006.10.023

Kubota, J. T., & Ito, T. (2017). Rapid race perception despite individuation and accuracy goals. *Social Neuroscience*, *12*(4), 468–478. https://doi.org/10.1080/17470919.2016.1182585

Kubota, J. T., Li, J., Bar-David, E., Banaji, M. R., & Phelps, E. A. (2013). The price of racial bias: Intergroup negotiations in the ultimatum game. *Psychological Science*, *24*(12), 2498–2504. https://doi.org/10.1177/0956797613496435

LaCosse, J., & Plant, E. A. (2020). Internal motivation to respond without prejudice fosters respectful responses in interracial interactions. *Journal of Personality and Social Psychology*, *119*(5), 1037-1056. https://doi.org/10.1037/pspi0000219

Lenth, R. V. (2022). Emmeans: Estimated marginal means, aka least-squares. R package version 1.7.2 https://CRAN.R-project.org/package=emmeans

LeVine, R. A., & Campbell, D. T. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior*. Oxford, UK: Wiley.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Maddox, K.B., Perry, J.M., & Pagan, J. (2022). Cues and categories: Revisiting paths to racial phenotypicality bias. *Social and Personality Psychology Compass, 16*, e12699. https://doi.org/10.1111/spc3.12699

Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology*, *41*(1), 19-35. https://doi.org/10.1016/j.jesp.2004.05.002

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*(1), 4–18. https://doi.org/10.1111/j.1469-8986.1995.tb03400.x

Martinez, E. & Kirchner, L. (2021, August 25). The secret bias hidden in mortgage-approval algorithms. *The Markup.* https://themarkup.org/denied/2021/08/25/the-secret-bias-hidden-in-mortgage-approval-algorithms

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Melson-Silimon, A., Spivey, B. N., & Skinner-Dorkenoo, A. L. (2024). The construction of racial stereotypes and how they serve as racial propaganda. *Social and Personality Psychology Compass, 18*(1), e12862. https://doi.org/10.1111/spc3.12862

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology, 65*(3), 469–485. https://doi.org/10.1037/0022-3514.65.3.469

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology*, *83*(5), 1029–1050. https://doi.org/10.1037/0022-3514.83.5.1029

Monteith, M. J., Lybarger, J. E., & Woodcock, A. (2009). Schooling the cognitive monster: The role of motivation in the regulation and control of prejudice. *Social and Personality Psychology Compass, 3*(3), 211–226. https://doi.org/10.1111/j.1751-9004.2009.00177.x

Murayama, K., Usami, S., & Sakaki, M. (2022). Summary-statistics-based power analysis: A new and practical method to determine sample size for mixed-effects modeling. *Psychological Methods, 27*(6), 1014–1038. https://doi.org/10.1037/met0000330

Navarrete, C. D., McDonald, M. M., Molina, L. E., & Sidanius, J. (2010). Prejudice at the nexus of race and gender: an outgroup male target hypothesis. *Journal of Personality and Social Psychology*, *98*(6), 933-945. https://doi.org/10.1037/a0017931

Page-Gould, E. (2017). Multilevel modeling. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *The Handbook of Psychophysiology* (4th ed., pp. 662-678). New York: Cambridge University Press.

Pauker, K., Lamer, S. A., Ansari, S., & Weisbuch, M. (2022). Cultural snapshots: Identifying cultural patterns that influence implicit racial bias. In K. C. McLean (Ed.), *Cultural*

methods in psychology: Describing and transforming cultures (pp. 109–145). Oxford University Press. https://doi.org/10.1093/oso/9780190095949.003.0004

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*(2), 181–192. https://doi.org/10.1037/0022-3514.81.2.181

Petsko, C. D., & Bodenhausen, G. V. (2020). Multifarious person perception: How social perceivers manage the complexity of intersectional targets. *Social and Personality Psychology Compass*, *14*(2), e12518.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*(5), 729–738. https://doi.org/10.1162/089892900562552

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811–832. https://doi.org/10.1037/0022-3514.75.3.811

Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(16), 6539–6544. https://doi.org/10.1073/pnas.0812589106

Rees, H. R., Ma, D. S., & Sherman, J. W. (2020). Examining the relationships among categorization, stereotype activation, and stereotype application. *Personality and Social Psychology Bulletin, 46*(4), 499-513. https://doi.org/10.1177/0146167219861431

Requin, J., Riehle, A., & Seal, J. (1988). Neuronal activity and information processing in motor control: from stages to continuous flow. *Biological Psychology*, *26*(1-3), 179–198. https://doi.org/10.1016/0301-0511(88)90019-1

Richeson, J. A., & Trawalter, S. (2008). The threat of appearing prejudiced & race-based attentional bias. *Psychological Science*, *19*, 98–102. https://doi.org/10.1111/j.1467-9280.2008.02052.x

Rugg, M. D., Milner, A. D., Lines, C. R., & Phalp, R. (1987). Modulation of visual event-related potentials by spatial and non-spatial visual selective attention. *Neuropsychologia*, *25*(1A), 85–96. https://doi.org/10.1016/0028-3932(87)90045-5

Salmon, C. A., & Hehman, J. A. (2022). The importance of ecological validity, ultimate causation, and natural categories. *The Behavioral and Brain Sciences*, *45*, e92. https://doi.org/10.1017/S0140525X21000741

Schmitt, H., Ferdinand, N.K. & Kray, J. (2015) The influence of monetary incentives on context processing in younger and older adults: an event-related potential study. *Cognitive, Affective, & Behavioral Neuroscience, 15*, 416–434. https://doi.org/10.3758/s13415-015-0335-x

Schuetz, J. (2020, December 9). *Rethinking homeownership incentives to improve household financial security and shrink the racial wealth gap*. Brookings. https://www.brookings.edu/research/rethinking-homeownership-incentives-to-improve-household-financial-security-and-shrink-the-racial-wealth-gap/

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, *23*(6), 695–703. https://doi.org/10.1111/j.1469-8986.1986.tb00696.x

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*(2), 314–335. https://doi.org/10.1037/0033-295X.115.2.314

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences*, *108*(19), 7710–7715. https://doi.org/10.1073/pnas.1014345108

Stanley, D. A., Sokol-Hessner, P., Fareri, D. S., Perino, M. T., Delgado, M. R., Banaji, M. R., & Phelps, E. A. (2012). Race and reputation: Perceived racial group trustworthiness influences the neural correlates of trust decisions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 744–753. https://doi.org/10.1098/rstb.2011.0300

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology, 35*(3), 413-424. https://doi.org/10.1002/ejsp.256

Todorov, A. (2010). The social perception of faces. In S.T. Fiske & C.N. Macrae (Eds.), *The Sage handbook of social cognition* (pp. 96-114). Thousand Oaks, CA: Sage.

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain Research*, *1494*, 44–55. https://doi.org/10.1016/j.brainres.2012.11.037

Tsotsos, J.K. (2005). The selective tuning model for visual attention. In L. Itti, G. Rees, & J.K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 562-569). New York: Academic Press.

Vaughan, A. C., & Birney, D. P. (2023). Within-individual variation in cognitive performance is not noise: Why and how cognitive assessments should examine within-person performance. *Journal of Intelligence*, *11*(6), 110. https://doi.org/10.3390/jintelligence11060110

Volpert-Esmond, H. I., & Bartholow, B. D. (2019). Explicit categorization goals affect attention-related processing of race and gender during person construal. *Journal of Experimental Social Psychology, 85,* Article 103839. https://doi.org/10.1016/j.jesp.2019.103839

Volpert-Esmond, H.I., & Bartholow, B.D. (2021). A functional coupling of brain and behavior during social categorization of faces. *Personality and Social Psychology Bulletin, 47*(11), 1580-1595. https://doi.org/10.1177/0146167220976688

Volpert-Esmond, H. I., Merkle, E. C., & Bartholow, B.D. (2017). The iterative nature of person construal: Evidence from event-related potentials. *Social Cognitive and Affective Neuroscience*, *12*(7), 1097-1107. https://doi.org/10.1093/scan/nsx048

Volpert-Esmond, H. I., Merkle, E. C., Levsen, M. P., Ito, T. A., & Bartholow, B. D. (2018). Using trial-level data and multilevel modeling to investigate within-task change in event-related potentials. *Psychophysiology*, *55*(5), e13044. https://doi.org/10.1111/psyp.13044

Volpert-Esmond, H. I., Page-Gould, E., & Bartholow, B. D. (2021). Using multilevel models for the analysis of event-related potentials. *International Journal of Psychophysiology*, *162*, 145–156. https://doi.org/10.1016/j.ijpsycho.2021.02.006

Von Gunten, C. D., Volpert-Esmond, H. I., & Bartholow, B. D. (2018). Temporal dynamics of reactive cognitive control as revealed by event-related brain potentials. *Psychophysiology*, *55*(3), 10.1111/psyp.13007. https://doi.org/10.1111/psyp.13007

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *326*(5960), 1711–1714. https://doi.org/10.1126/science.1178358

Willadsen-Jensen, E. C., & Ito, T. A. (2006). Ambiguity and the timecourse of racial perception. *Social Cognition, 24*(5), 580–606. https://doi.org/10.1521/soco.2006.24.5.580

Willadsen-Jensen, E., & Ito, T. A. (2015). The effect of context on responses to racially ambiguous faces: changes in perception and evaluation. *Social Cognitive and Affective Neuroscience*, *10*(7), 885–892. https://doi.org/10.1093/scan/nsu134

Zhang, T., Zhou, Y., & Han, S. (2023). Priority of racial and gender categorization of faces: A social task demand framework. *Journal of Personality and Social Psychology, 124*(3), 483–520. https://doi.org/10.1037/pspa0000318

Zhaoping, L. (2005). The primary visual cortex creates a bottom-up saliency map. In L. Itti, G. Rees, & J.K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 570-575). Academic Press.

Zhou, Y., Gao, T., Zhang, T., Li, W., Wu, T., Han, X., & Han, S. (2020). Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nature Human Behavior, 4*, 69–87. https://doi.org/10.1038/s41562-019-0743-y

Zilker, V., & Pachur, T. (2023). Attribute attention and option attention in risky choice. *Cognition*, *236*, 105441. https://doi.org/10.1016/j.cognition.2023.105441