Journal of Experimental Social Psychology 101 (2022) 104315

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Assessing the speed and spontaneity of racial bias in pain perception☆

Peter Mende-Siedlecki [a,*], Azaadeh Goharzad [a], Aizihaer Tuerxuntuoheti [a], Patrick Gilbert Mercado Reyes [a], Jingrun Lin [b], Alexis Drain [a]

[a] Department of Psychological & Brain Sciences, University of Delaware, United States
[b] Department of Psychology, University of Virginia, United States

ARTICLE INFO

Editor: Jack Rachael

Keywords:
Social perception
Pain
Racial bias
Emotion perception
Automaticity

ABSTRACT

A growing body of evidence demonstrates that perceivers recognize painful expressions less readily on Black (compared to White) faces. However, it is unclear how rapidly this bias emerges and whether it occurs automatically or effortfully—for example, via the deliberate application of consciously-held racialized beliefs regarding pain tolerance. Across five experiments we examined the speed and spontaneity of racial bias in pain perception. First, we observed that racial bias in pain perception was still evident under minimal presentation conditions (as brief as 33 ms) and was most apparent for ambiguous (versus high intensity) pain expressions (Exp. 1A-B). Notably, these findings generalized across both Black and White perceivers. Next, we manipulated the amount of cognitive load participants were under while viewing and rating Black and White faces in varying degrees of pain (Exps. 2A-C). Here, we observed that perceivers had more stringent thresholds for seeing pain on Black (versus White) faces regardless of whether participants were under high (versus low) load. Bayesian analyses of these data suggested strong evidence for the null hypothesis that racial bias in pain perception is not moderated by cognitive load. Together, these data demonstrate that racial bias in pain perception occurs rapidly, automatically, and with minimal visual input.

The pain of Black patients is consistently under-diagnosed and undertreated (Green et al., 2003). For example, Black patients are considerably less likely to receive opioids to treat both chronic non-cancer pain (Burgess et al., 2014; Burgess et al., 2014; Ringwalt, Roberts, Gugelmann, & Skinner, 2015) and cancer pain (Anderson et al., 2002), compared to White patients. Meta-analyses show that these gaps in care have remained in place for decades and extend across care contexts and types of pain (Lee et al., 2019; Meghani, Byun, & Gallagher, 2012). Since untreated pain has negative consequences for emotional and physical well-being (Katz, 2002; Niv & Kreitler, 2001; Wells, Pasero, & McCaffery, 2008), these gaps in pain care may have consequences for downstream disparities in disease morbidity, mortality, and disability affecting Black Americans (Fiscella & Sanders, 2016; Mays, Cochran, & Barnes, 2007). As such, a great deal of research has examined the psychological underpinnings of pain disparities.

Ultimately, long-standing disparities in pain care are a product of both historical and contemporary structural racism that has reduced Black Americans' access to healthcare and diminished the quality of that care (Feagin & Bennefield, 2014; Gee & Ford, 2011; Penner et al., 2013). Individual bias (e.g., at the level of providers) is rooted within that history and those structures. For example, while recent work implicates stereotypes regarding status, hardship, and false beliefs concerning biological differences between Black and White individuals (Deska et al., 2020; Hoffman, Trawalter, Axt, & Oliver, 2016; Trawalter, Hoffman, & Waytz, 2012; Waytz, Hoffman, & Trawalter, 2015) as key sources of biases in pain care, such stereotypes and beliefs are an outgrowth of racism in medical systems and beyond (Trawalter, Bart-Plange, & Hoffman, 2020). Indeed, the consequences of structural racism may even be so insidious as to manifest at the level of perception.

A growing body of work demonstrates that pain is recognized less readily on Black (compared to White) faces, stemming from disruptions in configural face processing (Mende-Siedlecki, Qu-Lee, Backer, & Van Bavel, 2019). This perceptual bias is comparatively stronger than race-based gaps in recognizing other emotions and has distinct consequences for treatment (Mende-Siedlecki et al., 2021). In addition, cues to racial prototypicality moderate this bias. For example, White perceivers' gaps in thresholds for seeing pain on Black (versus White) faces were magnified when targets were manipulated to have racially

☆ This paper has been recommended for acceptance by Dr Rachael Jack.

* Corresponding author at: Department of Psychological & Brain Sciences, University of Delaware, Wolf Hall, Newark DE 19716, United States.
E-mail address: pmende@udel.edu (P. Mende-Siedlecki).

https://doi.org/10.1016/j.jesp.2022.104315
Received 6 June 2021; Received in revised form 6 March 2022; Accepted 6 March 2022
Available online 5 April 2022

prototypic (versus neutral or non-prototypic) structural features (Drain, Goharzad, Qu-Lee, Lin, & Mende-Siedlecki, 2022). Similarly, this perceptual bias may vary as a function of target gender (Goharzad, Drain, Qu-Lee, Lin, & Mende-Siedlecki, 2022). Finally, meta-analysis across this work demonstrates that this perceptual bias generalizes across perceiver race (e.g., Black, Asian, and Hispanic White participants show the similar magnitudes of bias), stimuli (e.g., both photographic and computer-generated stimuli), and sample (e.g., student, online, and nursing samples; Lin, Drain, Goharzad, & Mende-Siedlecki, 2021). In addition, this tendency to see pain less readily on Black (versus White faces) is positively associated with a tendency to dehumanize Black (versus White) individuals and (at least within White perceivers) a tendency to live in more racially segregated communities. As we argue elsewhere (Lin et al., 2021), since contemporary racial dehumanization is an outgrowth of representations used to justify enslavement and oppression of Black people (Goff, Eberhardt, Williams, & Jackson, 2008) and since reduced intergroup contact within White individuals is a consequence of racial segregation and disparities in housing and lending (Pager & Shepherd, 2008; Pettigrew, 1998; Reskin, 2012), these data suggest that racial bias in pain perception may ultimately have deep roots in historical and structural racism.

While this evidence suggests that perceivers show differential sensitivity to pain on Black and White faces, our understanding of the mechanisms supporting this bias is limited. For example, it is unknown how rapidly racial bias in pain perception emerges, as well as whether this bias is influenced by the intensity (or ambiguity) of painful expressions. In addition, prior work has yet to ascertain whether this bias is automatic or dependent on effortful, deliberative processing. The present research comprises a series of investigations aimed at closing these gaps in knowledge.

1.1. Establishing the speed and spontaneity of racial bias in pain perception

Painful expressions—characterized by brow lowering, eyelid tightening, nose wrinkling, and upper lip and cheek raising (Kunz, Meixner, & Lautenbacher, 2019)—are robustly distinguished from other emotional expressions (Simon, Craig, Gosselin, Belin, & Rainville, 2008), processed rapidly in the visual system (Craig, Versloot, Goubert, Vervoort, & Crombez, 2010; Vervoort, Trost, Prkachin, & Mueller, 2013; Yamada & Decety, 2009), and their processing may be prioritized over other expressions at a neural level (González-Roldan et al., 2011; Reicherts et al., 2012). Thus, as far as a lower bound on the speed of racial bias in pain perception is concerned, it is possible that differences in sensitivity to pain on Black and White faces emerge with similar rapidity to the perception of painful expressions in general.

Rapid (versus unlimited) presentation increases the magnitude of racial biases in social and emotional perception. For example, while Black faces capture White perceivers' attention (Trawalter, Todd, Baird, & Richeson, 2008) and elicit enhanced amygdala activity relative to White faces when presented briefly (e.g., 30 ms), this difference is attenuated for longer presentation durations (e.g., 525 ms; Cunningham et al., 2004; Forbes, Cox, Schmader, & Ryan, 2012). In the context of emotion recognition, evidence for a happy categorization advantage for White faces and an angry categorization advantage for Black faces is observed at fast (200 ms) but not unlimited presentation durations (Craig, Mallan, & Lipp, 2012). As such, similar effects of presentation duration may govern racial bias in pain perception. However, if biased judgments of pain were *only* observed when perceivers have unlimited time to inspect stimuli, this might call into question the "perceptual" nature of this bias.

To address this question, we adapted an approach from researchers examining how quickly social evaluations are extracted from faces (e.g., Bar, Neta, & Linz, 2006; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). If a given facial characteristic can be recognized even under minimal presentation conditions, this suggests that processing pertaining to this characteristic is relatively automatic (Olson & Marshuetz, 2005). By that logic, if *bias* in recognition or evaluation is also present under minimal presentation conditions, this suggests that the bias is similarly automatic and not dependent on explicit attention or deliberation.

Another way to test the automaticity of this bias would be to hinder perceivers' ability to engage more deliberate processing mechanisms. Typical cognitive load manipulations tax working memory and deplete attentional resources, which, in turn, disrupts effortful processing (Gilbert & Osborne, 1989; Todorov & Uleman, 2003). Thus, if cognitive load reduces gaps between thresholds for seeing pain on Black and White faces, this might suggest that non-perceptual factors (e.g., explicit, consciously-held beliefs about racial differences in pain experience or pain tolerance) are primarily responsible for such effects. However, if bias in perceivers' recognition of pain on Black (versus White) faces is automatic, it should still be observed when perceivers are under cognitive load.

Initial evidence favors the automaticity account. For example, in a study conducted by Mathur, Richeson, Paice, Muzyka, and Chiao (2014), Black and White participants showed greater pro-White bias in their responses towards and treatment of the pain of White (versus Black) targets when patient race was primed implicitly, while this pattern reversed when patient race was presented explicitly. As such, those authors argue that prior work focusing specifically on explicit assessment of pain treatment or evaluation may underestimate the magnitude of racial bias in these measures. Supporting this possibility, in a meta-analysis comparing the strength of racial bias in pain perception to similar gaps in *treatment* decisions (a comparably explicit task), perceptual bias tended to be twice as large as treatment bias (Lin et al., 2021).

1.2. The present research

Across Experiments 1A-B and 2A-C, we adapted well-established paradigms to test the speed and spontaneity of racial bias in pain perception. Our investigation of the speed of this bias also offered an opportunity to examine whether its magnitude varies as a function of expression intensity. Our preregistered predictions were that this perceptual bias would be observed even under minimal presentation conditions and cognitive load. Taken together, this work adds to our understanding of the perceptual mechanisms supporting racial bias in pain perception.

2. Experiment 1A: assessing the speed of racial bias in pain perception

Painful expressions capture attention automatically and are processed rapidly (Craig et al., 2010; Vervoort et al., 2013; Yamada & Decety, 2009). However, race categorization exhibits a similarly rapid time course (Ito & Bartholow, 2009; Ito & Urland, 2003; Kubota & Ito, 2007) and this automatic extraction of racial category information occurs even when perceivers are attending to other social dimensions (e.g., gender; Ito & Urland, 2005) or motivated with individuation and accuracy goals (Kubota & Ito, 2017). While previous work demonstrates that perceivers consistently show more stringent thresholds for seeing pain on Black (versus White) faces, it is unclear if this bias is evident for brief presentations (reflecting automatic perceptual processes) or if this bias only occurs for longer presentations (reflecting more conscious, deliberate processes). Here, we examined how quickly racial bias in pain perception manifests by varying the presentation duration of Black and White targets making expressions of pain at varying intensity. We predicted that racial bias in pain perception would be evident under minimal presentation conditions (e.g., as brief as 50 ms). In addition, we predicted that both judgments of pain and racial bias within those judgments at brief presentation durations would be correlated with judgments and bias therein at unconstrained presentation time.

2.1. Methods

2.1.1. Participants

We recruited 132 Prolific participants living in the United States (61 men, 69 women, 2 non-binary; $M_{age} = 32.38$, $SD_{age} = 12.03$; 119 White, 12 Hispanic, and 1 participant who identified with another racial group[1]).

We preregistered our procedure, stimuli, sample size, and analysis plan (https://osf.io/8jczn). We sought a sample of 100 White participants. This sample size, while somewhat heuristically chosen, was more than twice the size of the largest sample in Todorov et al. (2009), which our design was modeled upon. Power analysis suggested that this size would afford us 80% power to detect an effect of $d = 0.285$ in a one-sample *t*-test (vs. 0, two-tailed; used to assess whether correlations between ratings made in each presentation speed bin are greater than chance). Given past spikes in fraudulent participation in online recruitment platforms, involving non-US participants using virtual private servers (Kennedy et al., 2020), we screened out some individuals prior to data collection with an established procedure for identifying participants using VPS/VPNs (Winter, Burleigh, Kennedy, & Clifford, 2019). Moreover, Prolific participants who completed similar previous studies were not able to access this task. Finally, at the beginning of the task, participants were asked to focus solely on the task while completing it (without other audio or visual distractions), and to complete the task indoors with the brightness on their computer screen set to 100% and the indoor lights dimmed if possible.

We also pre-registered and used a series of task-specific exclusion criteria. We removed participants from analysis who a) gave the same response on 90% of trials or greater, b) when asked if they completed the task alone answered "I was interrupted several times while completing the main task," c) when asked about their lighting and brightness answered "I completed the main task in a dimly-lit room but left my brightness level below 100%," "I adjusted my brightness level to 100%, but completed the main task either outdoors or in full lighting," or "I completed the main task either outdoors or in full lighting, and left my brightness level below 100%," d) when asked about their attention answered "I was actively watching or listening to something else while completing the main task," "I was working on other things while completing the main task," or "I had music or other audio on in the background while completing the main task," or e) gave an answer to an open-ended post-task question regarding the study's purpose that referred specifically to an interaction between target race and speed on thresholds for pain perception (confirmed if two out of three coders [the first, third, and fourth authors] determined that this rule had been violated).

These criteria resulted in the exclusion of 34 participants from analyses. The majority of these were individuals who did not make the requested changes to their lighting and screen brightness, in addition to two participants excluded based on their open-ended answers. However, we also excluded 10 participants who identified as White on Prolific but self-reported their race/ethnicity as another category in our post-task measures (9 Hispanic participants and 1 participant who identified with another racial group[2]). As such, our analyzed sample comprised 88 White Prolific participants living in the United States (43 male, 44 female, 1 non-binary; $M_{age} = 32.01$, $SD_{age} = 12.06$). (Thus, we had 80% power to detect an effect size of $d = 0.302$ in the key analysis described above, as per G*Power [v3.1].)

2.1.2. Stimuli

We selected 16 targets and four expressions from the digitally rendered subset of the Delaware Pain Database (DPD; Mende-Siedlecki, Qu-Lee, Lin, Drain, & Goharzad, 2020). In brief, to create these stimuli, we attempted to replicate painful expressions by manipulating sliders in FaceGen corresponding to facial action units, emotion expressions, and phonemes. The resulting expressions were normed in a characterization of the DPD; specifically, they were rated for their resemblance to pain and other emotions. Based on this norming data, the four pain expressions we selected were rated as looking more like pain on average than any other emotion ($M = 5.25$ out of 7; all comparison emotion $Ms < 3.26$, all comparison $ps < 0.0001$). Using these targets and expressions, we created 160 individual stimuli which varied in race (80 Black, 80 White) and pain intensity (each target/expression was created at five levels of intensity: 0%, 30%, 50%, 70% and 100%). Each expression was repeated four times within each race group. Target/expression pairs were equated across race such that each participant saw both Black and White versions of each pairing (Fig. 1).

2.1.3. Design and task

This experiment took the form of a 2 (Target race: Black vs. White) × 5 (Presentation speed: 33 ms, 50 ms, 100 ms, 500 ms, unconstrained) × 5 (Pain intensity: 0%, 30%, 50%, 70% and 100%) repeated measures design.[3] The task was constructed in PsychoPy (Peirce, 2007) and hosted on Pavlovia. Participants completed a modified paradigm adapted most directly from Experiment 2 in Todorov et al., 2009 (see also Willis and Todorov, 2006); they saw 160 male faces (80 Black, 80 White; created using FaceGen Pro) presented on a gray background, which varied in their expressions. Given the online format, we were unable to explicitly control visual angle. That said, we requested that participants maintain a viewing distance of 24 in. from their screens and we presented the stimuli such that each face was approximately 2.25 in. by 2.25 in. in size. As such, each face subtended approximately 5.4° both vertically and horizontally. Within the participants included in analyses, the modal screen resolution was 1366 × 768, with an average of 1581.87 × 929.49, based on metadata recorded from the post-task questionnaires via Qualtrics.

The 16 individual heads we selected were rendered them making the painful expressions described above at varying levels of intensity (0%, 30%, 50%, 70%, 100%). Black and White versions of each stimulus were created in FaceGen by varying the skin tone slider. Note that in FaceGen Pro, skin tone and structural prototypicality can be varied independently of each other. Given that structural cues to prototypicality can exacerbate racial bias in pain perception (Drain et al., 2022), in Experiments 1A and 1B, we chose to hold structure constant across Black and White stimuli for a more conservative test of our hypotheses. Thus, stimuli in these experiments had "average" structural features, in terms of the four racial groups accounted for in FaceGen's model of face space. Ultimately, stimuli were paired by race; a Black target making a given pain expression at the 70% level later appeared as a White target with identical facial structure making an identical expression.

Presentation speed was blocked across 5 different blocks. All 160 faces appeared in each block. Each trial began with a 500 ms fixation cross, followed by a stimulus – a Black or White face depicting some level of pain – with presentation length varying by block from 33 ms to unconstrained duration. This face was then replaced on screen with a Fourier-scrambled mask image. More specifically, for each target, we took an image of its base head with average skin tone (e.g., equidistant to the average Black and average White value in FaceGen, to avoid

[1] While we specifically sought to recruit White participants via Prolific, a small number of individuals identify as White in their Prolific profiles but self-reported their race/ethnicity as another category in our post-task measures.

[2] Results did not change meaningfully when these individuals were included in analyses (Supplementary Materials; https://osf.io/ht2u8).

[3] We also ran a previous pilot version of Experiment 1, the full procedure and results of which can be found in Supplementary Materials (https://osf.io/ht2u8). This pilot had an additional 67 ms condition but did not include 50% intensity morphs. The results of this pilot are strongly in accord with the findings of Experiment 1.

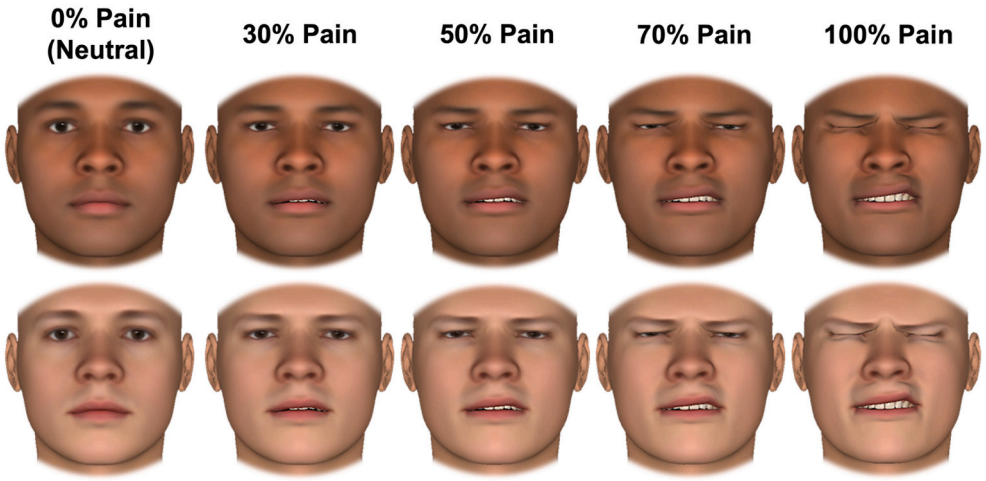

Fig. 1. Example stimuli from Experiments 1A-B. Participants saw morphs between neutral expressions and painful expressions rendered at five levels of intensity—0% (e.g., a neutral expression), 30%, 50%, 70%, and 100% pain. The Black and White faces depicted above a) were rendered from the same base head structure and b) are making the same painful expression.

differences in mask coloration across target race) and phase-scrambled that image using the function *imscramble* (Hebart, 2009). Below the perceptual mask, participants saw the question "Is this face in pain?" with a "Yes" or "No" response. Participants had unlimited time to make their responses with the Z and M keys. Mapping of the Z/M keys to Yes/No was counterbalanced across two versions of the task, which participants were randomly assigned to. A 1000 ms ITI (blank screen) separated each trial. Individual faces were shown in a randomized order within blocks and blocks (varying in presentation duration) were also shown in a randomized order.

Finally, participants were asked a series of questions: an open-ended item asking what they thought the study was about, a multiple-choice item assessing their strategy for assessing pain, and a series of multiple-choice questions related to our exclusion criteria.

2.1.3.1. Individual difference measures. Participants also completed exploratory individual difference measures of a) intergroup contact (brief version adapted from Cloutier, Li, & Correll, 2014), b) IMS/EMS (Plant & Devine, 1998), c) blatant dehumanization (Kteily, Bruneau, Waytz, & Cotterill, 2015), and d) explicit bias (e.g., feeling thermometers measuring warmth towards various social groups, including Black and White Americans). For this and all following experiments, analyses of these measures are collected in Supplementary Materials available online.

2.1.4. Analyses

We first recoded all "Yes" and "No" judgments as 1 s and 0 s, respectively, and calculated average responses within each cell of the 2 × 5 × 5 repeated measures design. Higher numbers within a given cell reflect more stimuli being judged as looking like they were in pain.

We preregistered a predicted main effect of race on pain judgments; participants would be more likely to judge White (versus Black) targets as being in pain overall. To test this, we performed a 2 (target race) × 5 (presentation length) × 5 (pain intensity) repeated measures ANOVA. While we did not preregister a prediction regarding the effect of presentation length (or its interaction with target race), we did generally expect that bias would be larger in the ambiguous intensities (e.g., 30–70%), compared to the 0% or 100% intensity bins.

We also predicted that positive correlations between a) speeded and unconstrained judgments of pain and b) *racial bias* in speeded and unconstrained judgments of pain would be detectable above chance by 50 ms. To calculate bias within each pair of targets, we simply subtracted the Black pain judgment from the White pain judgment for each pair (since each Black target has a corresponding White target—an identical face, in terms of structure, making the same expression). Thus, if a participant judged the White version of a given pair as being in pain ("Yes" = 1), but judged the Black version of that pair as *not* being in pain ("No" = 0), then this participant's bias value for the pair in question would be 1 minus 0, or 1. Higher numbers within a given cell reflect greater racial bias in pain judgments.

To assess both correlational predictions, we followed the approach of Todorov et al. (2009) and first calculated bivariate correlations between speeded and unconstrained judgments of pain averaged across participants at the level of individual stimuli. While we pre-registered this analytic strategy, we later felt it was necessary to also demonstrate that these relationships were not merely due to clustering of responses as a function of pain intensity. As such, we also present partial correlations controlling for pain intensity.

As Todorov et al. (2009) noted when conducting a similar analysis on speeded judgments of trustworthiness, these correlations can be inflated by aggregating judgments across participants at the stimulus level. Thus, we conducted a second analysis at the level of individual participants. For each participant and at each level of presentation duration, we computed correlations between the participant's speeded pain judgments and the mean criterion ratings (e.g., averaged across all participants in the unconstrained condition). We computed a similar set of correlations for racial *bias* in pain judgments. We then transformed these raw correlations into Fisher *z*-scores that could be used in subsequent statistical analyses.

Our procedure for determining sample size, all data exclusions, all manipulations, and all measures included in this research are fully reported in this article. All stimuli materials and de-identified data for this and all experiments have been made available online (https://osf.io/ht2u8). Further, Supplementary Materials are available online which offer additional details on each experiment—including analyses of results including all individuals passing task-based exclusion criteria (e.g., regardless of participant race/ethnicity) and analyses of correlational data related to the exploratory individual difference measures we collected.

2.2. Results

2.2.1. Effects of target race, presentation duration, and pain intensity on pain judgments

We observed statistically significant main effects of target race ($F(1,87) = 64.23$, $p < .001$, $\eta_p^2 = 0.43$), presentation duration ($F(4,348) =$

2.44, $p = .047$, $\eta_p^2 = 0.03$), and pain intensity ($F(4,348) = 743.16$, $p < .001$, $\eta_p^2 = 0.90$) on pain judgments. As predicted, participants were more likely to judge White versus Black faces as being in pain ($M_{Black} = 0.526$, $SD_{Black} = 0.129$; $M_{White} = 0.551$, $SD_{White} = 0.123$). Moreover, as expected, participants' pain judgments tracked linearly with the pain intensity of targets' expressions ($M_{0\%} = 0.051$, $SD_{0\%} = 0.064$; $M_{30\%} = 0.246$, $SD_{30\%} = 0.223$; $M_{50\%} = 0.571$, $SD_{50\%} = 0.269$; $M_{70\%} = 0.869$, $SD_{70\%} = 0.139$; $M_{100\%} = 0.954$, $SD_{100\%} = 0.067$). Interestingly, pain judgments were highest within the 100 ms presentation bin ($M_{100ms} = 0.557$, $SD_{100ms} = 0.145$), and these ratings were significantly different from three of the other four presentation bins ($M_{33ms} = 0.529$, $SD_{33ms} = 0.148$; $M_{500ms} = 0.534$, $SD_{500ms} = 0.132$; $M_{unconstrained} = 0.532$, $SD_{unconstrained} = 0.128$), but not the 50 ms presentation bin ($M_{50ms} = 0.542$, $SD_{50ms} = 0.139$).

Notably, we did *not* observe a significant two-way interaction between target race and presentation duration [$F(4,348) = 1.41$, $p = .230$, $\eta_p^2 = 0.02$]. Participants were more likely to report seeing pain on White (versus Black) faces within the 33 ms ($MD = 0.026$, $SE = 0.006$, $p < .001$, $\eta_p^2 = 0.20$), 50 ms ($MD = 0.022$, $SE = 0.006$, $p < .001$, $\eta_p^2 = 0.14$), 100 ms ($MD = 0.018$, $SE = 0.005$, $p = .001$, $\eta_p^2 = 0.13$), 500 ms ($MD = 0.032$, $SE = 0.005$, $p < .001$, $\eta_p^2 = 0.32$), and unlimited presentation bins ($MD = 0.031$, $SE = 0.005$, $p < .001$, $\eta_p^2 = 0.28$).

However, we did observe a two-way interaction between target race and pain intensity ($F(4,348) = 18.04$, $p < .001$, $\eta_p^2 = 0.17$; Fig. 2). In line with our general predictions, racial bias in pain judgments (e.g., a tendency to judge more White targets as being in pain versus Black targets) was largest for the 50% pain intensity expressions ($F(1,87) = 51.12$, $p < .001$, $\eta_p^2 = 0.37$; $M_{Black} = 0.541$, $SD_{Black} = 0.280$; $M_{White} = 0.601$, $SD_{White} = 0.263$), somewhat smaller for the 30% pain intensity ($F(1,87) = 26.43$, $p < .001$, $\eta_p^2 = 0.23$; $M_{Black} = 0.227$, $SD_{Black} = 0.224$; $M_{White} = 0.266$, $SD_{White} = 0.227$) and 70% pain intensity expressions ($F(1,87) = 17.33$, $p < .001$, $\eta_p^2 = 0.17$; $M_{Black} = 0.856$, $SD_{Black} = 0.150$; $M_{White} = 0.881$, $SD_{White} = 0.134$), and only marginally significant for the 0% pain intensity expressions ($F(1,87) = 3.74$, $p = .056$, $\eta_p^2 = 0.04$; $M_{Black} = 0.047$, $SD_{Black} = 0.060$; $M_{White} = 0.055$, $SD_{White} = 0.073$). This effect was not statistically significant within the 100% pain intensity expressions ($F(1,87) = 0.99$, $p = .323$, $\eta_p^2 = 0.01$; $M_{Black} = 0.956$, $SD_{Black} = 0.063$; $M_{White} = 0.953$, $SD_{White} = 0.073$).[4]

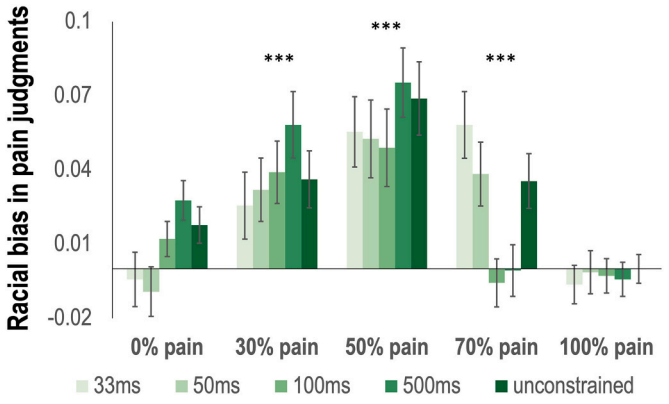

Fig. 2. Racial bias in pain recognition is rapid and moderated by ambiguity. Participants in Experiment 1A saw pain less readily on Black (versus White) faces even when presentation duration was as low as 33 ms (lightest bars). Moreover, target race interacted with expression intensity; bias was largest when stimuli were most ambiguous (e.g., the 50% morphs). Error bars represent within-subjects corrected SEM. (***$p < .001$; *$p < .05$; for a box-and-whisker plot collapsed across presentation duration with individual data points overlaid, see Supplementary Materials).

2.2.2. Effects of presentation duration on correspondence between speeded and unconstrained judgments

Pain judgments aggregated at the stimulus level were already strongly positively correlated with unconstrained judgments by 33 ms ($r = 0.986$, $p < .001$). In other words, face stimuli that were judged as expressing pain after 33 ms presentation were more likely to be judged as expressing pain when stimuli were presented for an unlimited period. These strong, positive correlations held for presentation durations of 50 ms ($r = 0.987$, $p < .001$), 100 ms ($r = 0.990$, $p < .001$), and 500 ms ($r = 0.995$, $p < .001$), and further, when we controlled for the objective degree of pain intensity in each expression (33 ms: $r = 0.843$, $p < .001$; 50 ms: $r = 0.867$, $p < .001$; 100 ms: $r = 0.893$, $p < .001$; 500 ms: $r = 0.948$, $p < .001$).

As described above, we also conducted a second analysis at the level of individual participants. The average Fisher-transformed correlation between individual participants' judgments and the mean criterion ratings was 0.897 ($SD = 0.250$) for 33 ms presentation durations. This correlation stayed strong and positive for presentation durations of 50 ms (avg. Fisher-transformed $r = 0.912$, $SD = 0.248$), 100 ms (avg. Fisher-transformed $r = 0.951$, $SD = 0.217$), and 500 ms (avg. Fisher-transformed $r = 1.018$, $SD = 0.225$). In each case, the average correlation observed was significantly different from zero in a one-sample t-test (33 ms: $t(87) = 33.60$, $p < .001$; 50 ms: $t(87) = 34.50$, $p < .001$; 100 ms: $t(87) = 41.09$, $p < .001$; 500 ms: $t(87) = 42.35$, $p < .001$). Taken together, pain content is rapidly extracted from facial stimuli.

2.2.3. Effects of presentation duration on correspondence between racial bias in speeded and unconstrained judgments

Moreover, racial bias in pain judgments aggregated at the stimulus level was also already positively correlated with unconstrained judgments by 33 ms ($r = 0.389$, $p < .001$). In other words, a tendency to see more pain on a White (versus a Black) version of a given stimulus at 33 ms presentation was positively associated with the same racial bias in pain judgments when that stimulus was presented for an unlimited period of time. These strong, positive correlations held for presentation durations of 50 ms ($r = 0.495$, $p < .001$), 100 ms ($r = 0.360$, $p = .001$), and 500 ms ($r = 0.472$, $p < .001$). Each of these correlations was conserved when controlling for the objective degree of pain intensity in each expression (33 ms: $r = 0.399$, $p < .001$; 50 ms: $r = 0.506$, $p < .001$; 100 ms: $r = 0.348$, $p = .002$; 500 ms: $r = 0.462$, $p < .001$).

Assessing these effects at the level of individual participants weakened these associations (33 ms: avg. Fisher-transformed $r = 0.061$, $SD = 0.126$; 50 ms: avg. Fisher-transformed $r = 0.080$, $SD = 0.131$; 100 ms: avg. Fisher-transformed $r = 0.057$, $SD = 0.134$; 500 ms: avg. Fisher-transformed $r = 0.064$, $SD = 0.136$). That said, the average correlation between individual-level bias and mean criterion ratings was still significantly different from zero within each presentation duration (33 ms: $t(87) = 4.50$, $p < .001$; 50 ms: $t(87) = 5.72$, $p < .001$; 100 ms: $t(87) = 4.00$, $p < .001$; 500 ms: $t(87) = 4.38$, $p < .001$; one-sample t-tests versus zero). Taken together, not only is racial bias in pain judgments evident after minimal presentation times, but bias evident at 33 ms is positively correlated with bias in unconstrained judgments.

2.3. Discussion

Experiment 1A demonstrated how rapidly racial bias in pain

[4] We also observed a two-way interaction between presentation duration and intensity ($F(16,1392) = 5.95$, $p < .001$, $\eta_p^2 = 0.06$), as well as a three-way interaction between target race, presentation duration, and pain intensity ($F(16,1392) = 2.24$, $p = .003$, $\eta_p^2 = 0.03$; see Supplementary Materials).

perception emerges. Participants displayed a tendency to see pain more readily on White versus Black faces after only 33 ms presentation. While this bias was observed across all presentation durations (including unconstrained presentation), it was largest in magnitude for the most ambiguous expressions of pain (e.g., the 50% morphs).[5] This finding is in step with other work suggesting that group-based perceptual biases are magnified under conditions of ambiguity (Freeman, Penner, Saperstein, Scheutz, & Ambady, 2011; Xiao, Coppin, & Van Bavel, 2016). Critically, not only was bias in pain perception quick to emerge, but bias observed as early as 33 ms was positively correlated with bias at unconstrained presentation durations. Taken together, these data support the conclusion that racial bias in pain perception occurs even under minimal presentation conditions.

3. Experiment 1B: confirming the speed of racial bias in pain perception

Next, we sought to test the robustness of the effects observed in Experiment 1A. Moreover, we examined whether the perceptual bias demonstrated in Experiment 1A (as well as its moderation by expression intensity) would generalize to Black perceivers. Meta-analysis across prior research on racial bias in pain perception suggests that Black perceivers also demonstrate a tendency to see pain less readily on Black (versus White) faces (Lin et al., 2021), implying that this perceptual bias does not merely reflect general in-group favoritism. These findings are in step with other work suggesting that Black participants show similar biases in pain attribution to White participants (e.g., Deska et al., 2020; Trawalter et al., 2012). That said, we have yet to explicitly compare between Black and White participants in the same investigation of racial bias in pain perception.

While Experiment 1B was largely a direct replication of Experiment 1A, we added a subset of non-painful (e.g., angry and happy) expressions and informed participants of their presence in the stimuli set, so as to focus their attention specifically on pain. Lastly, we used visual masks that were individually tailored to the faces they were masking.

3.1. Methods

3.1.1. Participants

We recruited 270 Prolific participants living in the United States (106 men, 154 women, 9 non-binary, 1 preferred to self-describe; M_{age} = 31.41, SD_{age} = 11.20; 141 White, 127 African American, 2 preferred to self-describe; 265 non-Hispanic/Latinx, 5 Hispanic/Latinx).

We preregistered our procedure, stimuli, sample size, and analysis plan (https://osf.io/xfyjv). Our sample size was consistent with Experiment 1A. Given our aim of testing the generalizability of those findings across perceiver race, we aimed for 100 Black and 100 White participants. We implemented the same instructions regarding attention and environmental distractions as in Experiment 1A, as well as the same screening procedures and requested viewing distance. These criteria resulted in the exclusion of 75 participants from analyses. The majority of these were individuals who did not make the requested changes to their lighting and screen brightness, but also included 1) 3 participants passing screening who identified as White or Black on Prolific but self-reported their race/ethnicity as another category in our post-task measures, 2) 3 White participants passing screening who enrolled after the 100th White participant finished the task, 3) 15 participants who didn't finish the main task but completed the post-task, and 4) 8 individuals who completed the main task but didn't advance to the post-task (and as such, are not included in the demographics presented above). Ultimately, our analyzed sample comprised 100 White and 95 Black Prolific participants living in the United States (109 female, 76 male, 10 non-binary; M_{age} = 31.04, SD_{age} = 11.42). Thus, considering these two groups separately, we had 80% power to detect an effect size of at least d = 0.292 [within the Black participants, specifically] in the key analysis described above, as per G*Power [v3.1].

3.1.2. Stimuli

In addition to the stimuli used in Experiment 1A, we also included additional stimuli depicting high-intensity angry and happy expressions. Specifically, 8 happy Black faces, 8 happy White faces, 8 angry Black faces, and 8 angry White faces were included. These expressions were rendered at 100% intensity and randomly intermixed with the other 160 painful stimuli. The inclusion of angry and happy expressions served a dual purpose. First, by informing participants that other (e.g., non-pain) expressions would appear, we ensured that participants were more specifically focused on detecting pain-specific content in facial expressions. Second, this approach allowed us to determine if participants were extracting specific emotion content from the expressions they saw during the task, rather than merely responding to *any* change in facial expression that they saw. All things being equal, participants should be more likely to give a "Yes" response to a pain expression, compared to an anger or a happy expression. Including these stimuli allowed us to test whether this was the case, and if so, how quickly participants were able to make this distinction. Ultimately, for racial bias in pain perception to occur rapidly, it follows that pain detection *overall* should be similarly rapid.

These expressions were rendered on base heads already included in the stimulus set. Based on previous pilot testing of these expressions (Mende-Siedlecki et al., 2021), the four angry expressions we used were rated, on average, as looking more like anger (M = 5.20, SD = 0.50) than pain (M = 2.65, SD = 0.59, p = .002) or happiness (M = 1.66, SD = 0.17, p < .001), while the four happy expressions we used were rated, on average, as looking more like happiness (M = 5.25, SD = 0.39) than pain (M = 2.03, SD = 0.12, p < .001) or anger (M = 2.01, SD = 0.14, p < .001). Moreover, the selected anger and happiness expressions were rated as looking more like anger and happiness respectively, than the selected pain expressions (M_{Anger} = 2.99, SD_{Anger} = 0.50, p < . 001; M_{Happy} = 2.17, SD_{Happy} = 0.39, p < .001), while the selected pain expressions were rated as looking more like pain (M_{Pain} = 5.27, SD_{Pain} = 0.11) than either the selected anger (p < .001) or happiness expressions (p < .001).

We also made a slight change to our mask stimuli. In Experiment 1A, we made masks from average skin tone versions of each base head to avoid differences in mask coloration across target race. However, to rule out possibility that the resulting masks were more effective for masking one level of target race than the other, masks in Experiment 1B were made *directly from* the targets they were masking.

3.1.3. Design and task

Our task procedures and display parameters were the same as in Experiment 1A, with the addition of the 32 angry and happy stimuli and the switch to target-specific masks. Within the participants included in analyses, the modal screen resolution was 1366 × 768, with an average of 1476.61 × 863.19.

3.1.4. Analyses

Our basic analytic approach and prediction were very similar to those described for Experiment 1A. However, we note several additions. First, we now pre-registered a prediction regarding the target race by intensity interaction observed in Experiment 1A. In brief, we predicted

[5] These results replicate those of Supplementary Experiment 1 (a previous pilot version of Experiment 1A; results and materials available online at https://osf.io/ht2u8/?view_only=8261e33dfb7a4521be3242211a788556), lending additional confidence to our observations. Participants in this previous experiment also saw pain less readily on Black (versus White) target faces, even when stimuli were presented as briefly as 33 ms. Moreover, target race also interacted with presentation duration; this perceptual bias was observed for 30% and 70% intensity pain expressions, but not 0% and 100% intensity stimuli. (Supplementary Experiment 1 did not employ a 50% intensity condition.)

that racial bias in pain judgments would be largest for the most ambiguous expressions of pain and would decrease as expressions became both more and less intense. Second, we also predicted that all primary effects (e.g., the main effect of target race, the target race by intensity interaction, the correlations between judgments of pain and bias therein at speeded and unconstrained presentation, and the robustness of these effects across presentation duration) would generalize across *perceiver* race. Finally, we predicted that participants would be more likely to judge 100% pain expressions as looking like pain, compared to the 100% anger and 100% happy faces now included in the task, and further, that this effect would hold across presentation duration and perceiver race.

3.2. Results

3.2.1. Expression specificity manipulation check

We predicted that participants' judgments of pain would be sensitive to the actual content of expressions, as opposed to the mere presence of *any* expression on a given face. As such, participants should be more likely to give a "Yes" response to a painful expression, compared to an angry or a happy expression. Indeed, we observed a main effect of expression type on pain judgments of 100% intensity expressions varying in emotion between pain, anger, and happiness ($F(2,388) = 990.31$; $p < .001$, $\eta_p^2 = 0.84$). Post-hoc comparisons revealed that participants were significantly more likely to judge 100% pain expressions as looking like pain ($M = 0.874$, $SD = 0.139$) compared to 100% angry ($M = 0.793$, $SD = 0.217$; $p < .001$) and 100% happy expressions ($M = 0.143$, $SD = 0.174$; $p < .001$).

Critically, and as predicted, this effect was observed at each level of presentation duration (33 ms: $F(2,388) = 426.61$; $p < .001$, $\eta_p^2 = 0.69$; 50 ms: $F(2,388) = 536.09$; $p < .001$, $\eta_p^2 = 0.73$; 100 ms: $F(2,388) = 839.82$; $p < .001$, $\eta_p^2 = 0.81$; 500 ms: $F(2, 388) = 1062.75$; $p < .001$, $\eta_p^2 = 0.85$; unconstrained: $F(2,388) = 1088.60$; $p < .001$, $\eta_p^2 = 0.85$). (Though the difference in judgments was clearly larger between painful and happy expressions, the pain vs. anger comparison was statistically significant at each level of presentation duration [all ps < 0.001].) In addition, this effect was observed within both Black ($F(2,188) = 696.47$; $p < .001$, $\eta_p^2 = 0.88$) and White perceivers ($F(2,388) = 394.31$; $p < .001$, $\eta_p^2 = 0.80$) as predicted, and the interaction between emotion and perceiver race was not statistically significant ($F(2,386) = 2.06$; $p = .129$, $\eta_p^2 = 0.01$). Taken together, these results suggest that participants were extracting specific information about the content of emotional expressions (even after only 33 ms presentation), and not simply making their judgments based on general changes in facial configurations.

3.2.2. Effects of target race, presentation duration, pain intensity, and participant race on pain judgments

In addressing our primary hypotheses, we begin by presenting results that collapse across participant race. Across both Black and White perceivers, we observed statistically significant main effects of target race ($F(1,193) = 129.26$; $p < .001$, $\eta_p^2 = 0.40$), presentation duration ($F(4,772) = 7.94$; $p < .001$, $\eta_p^2 = 0.04$), and pain intensity ($F(4,772) = 986.43$; $p < .001$, $\eta_p^2 = 0.84$) on pain judgments. As predicted, participants were more likely to judge White versus Black faces as being in pain ($M_{Black} = 0.460$, $SD_{Black} = 0.112$; $M_{White} = 0.492$, $SD_{White} = 0.110$).

Moreover, as expected, participants' pain judgments tracked linearly with the pain intensity of targets' expressions ($M_{0\%} = 0.109$, $SD_{0\%} = 0.164$; $M_{30\%} = 0.184$, $SD_{30\%} = 0.182$; $M_{50\%} = 0.449$, $SD_{50\%} = 0.199$; $M_{70\%} = 0.763$, $SD_{70\%} = 0.179$; $M_{100\%} = 0.874$, $SD_{100\%} = 0.139$). In general, pain judgments increased as presentation durations grew longer ($M_{33ms} = 0.456$, $SD_{33ms} = 0.149$; $M_{50ms} = 0.462$, $SD_{50ms} = 0.138$; $M_{100ms} = 0.487$, $SD_{100ms} = 0.126$; $M_{500ms} = 0.496$, $SD_{500ms} = 0.123$; $M_{unconstrained} = 0.479$, $SD_{unconstrained} = 0.122$). Here, the only durations for which judgments weren't significantly different from each other were the 33 ms and the 50 ms durations, the 100 ms and the 500 ms durations, and the 100 ms and the unconstrained durations.

Once again, we did not observe a significant two-way interaction between target race and presentation duration ($F(4,772) = 0.278$; $p = .892$, $\eta_p^2 = 0.001$). In other words, the effect of target race generalized across duration. Participants were more likely to report seeing pain on White (versus Black) faces within the 33 ms ($F(1,193) = 45.88$; $p < .001$, $\eta_p^2 = 0.192$; $M_{Black} = 0.439$, $SD_{Black} = 0.155$; $M_{White} = 0.473$, $SD_{White} = 0.152$), 50 ms ($F(1,193) = 38.44$; $p < .001$, $\eta_p^2 = 0.166$; $M_{Black} = 0.447$, $SD_{Black} = 0.141$; $M_{White} = 0.477$, $SD_{White} = 0.145$), 100 ms ($F(1,193) = 56.87$; $p < .001$, $\eta_p^2 = 0.228$; $M_{Black} = 0.470$, $SD_{Black} = 0.133$; $M_{White} = 0.503$, $SD_{White} = 0.126$), 500 ms ($F(1,193) = 50.98$; $p < .001$, $\eta_p^2 = 0.209$; $M_{Black} = 0.480$, $SD_{Black} = 0.123$; $M_{White} = 0.511$, $SD_{White} = 0.130$), and unlimited presentation bins ($F(1,193) = 64.32$; $p < .001$, $\eta_p^2 = 0.250$; $M_{Black} = 0.461$, $SD_{Black} = 0.122$; $M_{White} = 0.496$, $SD_{White} = 0.129$).

As in Experiment 1A and in accordance with our pre-registered predictions, we also observed a significant two-way interaction between target race and pain intensity ($F(4,772) = 40.68$; $p < .001$, $\eta_p^2 = 0.174$). As predicted, racial bias in pain perception was largest for the most ambiguous expressions of pain: the 50% pain intensity expressions $F(1,194) = 116.87$; $p < .001$, $\eta_p^2 = 0.376$; $M_{Black} = 0.411$, $SD_{Black} = 0.208$; $M_{White} = 0.487$, $SD_{White} = 0.203$). This bias was smaller but still statistically significant for 30% ($F(1,194) = 68.95$; $p < .001$, $\eta_p^2 = 0.262$; $M_{Black} = 0.163$, $SD_{Black} = 0.186$; $M_{White} = 0.204$, $SD_{White} = 0.186$) and 70% expressions ($F(1,194) = 38.33$; $p < .001$, $\eta_p^2 = 0.165$; $M_{Black} = 0.747$, $SD_{Black} = 0.188$; $M_{White} = 0.779$, $SD_{White} = 0.177$), and even smaller for 0% expressions ($F(1,194) = 7.93$; $p = .005$, $\eta_p^2 = 0.039$; $M_{Black} = 0.104$, $SD_{Black} = 0.162$; $M_{White} = 0.115$, $SD_{White} = 0.170$). This bias was not statistically significant within 100% pain intensity expressions ($F(1,194) = 0.244$; $p = .622$, $\eta_p^2 = 0.001$; $M_{Black} = 0.873$, $SD_{Black} = 0.138$; $M_{White} = 0.875$, $SD_{White} = 0.142$). Moreover, this same interaction between target race and pain intensity was observed at each level of presentation duration (33 ms: $F(4,772) = 6.495$; $p < .001$, $\eta_p^2 = 0.033$; 50 ms: $F(4, 772) = 13.435$; $p < .001$, $\eta_p^2 = 0.065$; 100 ms: $F(4,772) = 8.689$; $p < .001$, $\eta_p^2 = 0.043$; 500 ms: $F(4,772) = 14.726$; $p < .001$, $\eta_p^2 = 0.071$; unconstrained presentation: $F(4,772) = 24.919$; $p < .001$, $\eta_p^2 = 0.114$).

3.2.2.1. Interactions with perceiver race. The results above collapse across Black and White perceivers. That said, our pre-registered predictions stated that our primary hypotheses would be confirmed *within* both Black and White perceivers when considered separately.

The main effect of target race on pain judgments described above was indeed observed in both Black ($F(1,94) = 100.12$; $p < .001$, $\eta_p^2 = 0.516$; $M_{BlackFaces} = 0.445$, $SD_{BlackFaces} = 0.106$; $M_{WhiteFaces} = 0.488$, $SD_{WhiteFaces} = 0.108$) and White perceivers ($F(1,99) = 34.97$; $p < .001$, $\eta_p^2 = 0.261$; $M_{BlackFaces} = 0.496$, $SD_{BlackFaces} = 0.113$; $M_{WhiteFaces} = 0.473$, $SD_{WhiteFaces} = 0.117$). That said, we do note that we also observed a target race by perceiver race interaction ($F(4,772) = 10.96$; $p = .001$, $\eta_p^2 = 0.054$), such that this tendency to judge expressions as looking like pain more readily on White versus Black faces was actually even larger within Black perceivers. Furthermore, racial bias in pain judgments was observed at each level of presentation duration for both Black and White perceivers (all ps < 0.027; observed for White perceivers in the 50 ms condition).

The interaction between target race and pain intensity was observed within both Black ($F(4,376) = 21.82$; $p < .001$, $\eta_p^2 = 0.188$) and White perceivers ($F(4,376) = 19.14$; $p < .001$, $\eta_p^2 = 0.162$; Fig. 3). In both cases, this interaction took the same pattern as described above: bias in pain judgments was largest for the most ambiguous (e.g., 50%) pain expressions, and decreased as expressions became both more neutral and more extreme. Moreover, we did not observe an interaction between target race and perceiver race ($F(4,772) = 0.63$; $p = .625$, $\eta_p^2 = 0.003$).

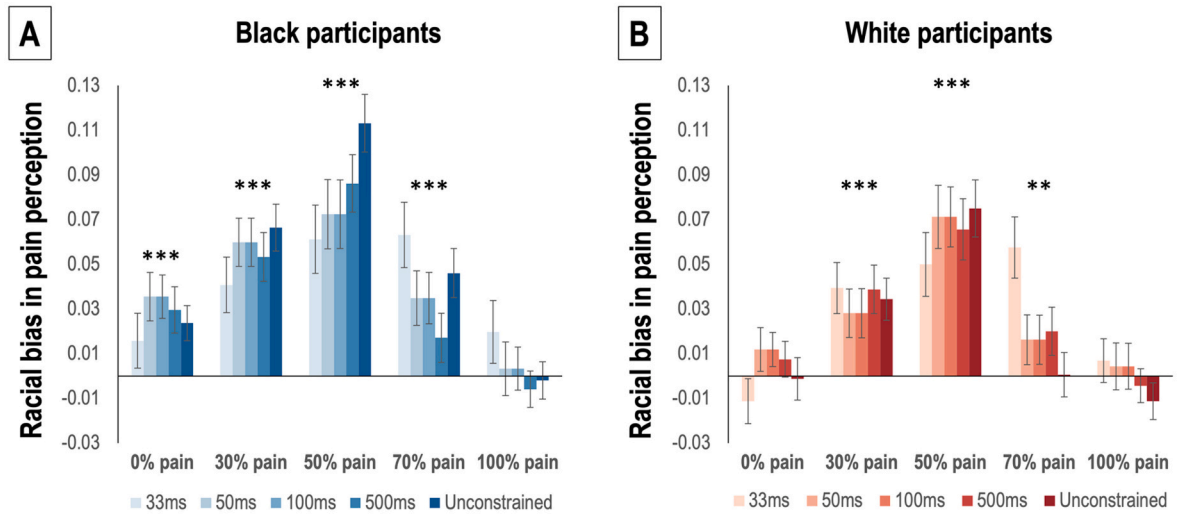

Fig. 3. Rapid racial bias in pain recognition generalizes across perceiver race. Both Black participants (A) and White participants (B) in Experiment 2 saw pain less readily on Black (versus White) faces. Once again, this bias was evident for presentations as fast as 33 ms and bias was still largest when expression intensity was most ambiguous (e.g., the 50% morphs). Error bars represent within-subjects corrected SEM. (***$p < .001$; **$p < .05$; *$p < .05$; for a box-and-whisker plot collapsed across presentation duration with individual data points overlaid, see Supplementary Materials).

3.2.3. Effects of presentation duration on correspondence between speeded and unconstrained judgments

Pain judgments aggregated at the stimulus level were already strongly positively correlated with unconstrained judgments by 33 ms ($r = 0.984$, $p < .001$). In other words, face stimuli that were judged as expressing pain after 33 ms presentation were more likely to be judged as expressing pain when stimuli were presented for an unlimited period. These strong, positive correlations held for presentation durations of 50 ms ($r = 0.991$, $p < .001$), 100 ms ($r = 0.993$, $p < .001$), and 500 ms ($r = 0.996$, $p < .001$), and further, when we controlled for the objective degree of pain intensity in each expression (33 ms: $r = 0.864$, $p < .001$; 50 ms: $r = 0.920$, $p < .001$; 100 ms: $r = 0.938$, $p < .001$; 500 ms: $r = 0.961$, $p < .001$).

As described above, we also conducted a second analysis at the level of individual participants. The average Fisher-transformed correlation between individual participants' judgments and the mean criterion ratings was 0.512 ($SD = 0.328$) for 33 ms presentation durations. This correlation stayed strong and positive for presentation durations of 50 ms (avg. Fisher-transformed $r = 0.605$, $SD = 0.343$), 100 ms (avg. Fisher-transformed $r = 0.676$, $SD = 0.309$), and 500 ms (avg. Fisher-transformed $r = 0.772$, $SD = 0.329$). In each case, the average correlation observed was significantly different from zero in a one-sample *t*-test (33 ms: $t(194) = 21.79$, $p < .001$; 50 ms: $t(194) = 24.66$, $p < .001$; 100 ms: $t(194) = 30.52$, $p < .001$; 500 ms: $t(194) = 32.82$, $p < .001$). Taken together, pain content is rapidly extracted from facial stimuli.

3.2.3.1. Interactions with perceiver race. Each of the stimulus-level correlations described above generalized across perceiver race, both in terms of the zero-order bivariate correlations (all $rs > 0.973$, all $ps < 0.001$) and the partial correlations controlling for expression intensity (all $rs > 0.804$, all $ps < 0.001$). When focusing on the individual-level correlations, these results also generalized across perceiver race (all average Fisher-transformed $rs > 0.465$, all $ps < 0.001$).

3.2.4. Effects of presentation duration on correspondence between racial bias in speeded and unconstrained judgments

Moreover, racial bias in pain judgments aggregated at the stimulus level was also already positively correlated with unconstrained judgments by 33 ms ($r = 0.418$, $p < .001$). In other words, a tendency to see more pain on a White (versus a Black) version of a given stimulus at 33 ms presentation was positively associated with the same racial bias in pain judgments when that stimulus was presented for an unlimited period of time. These strong, positive correlations held for presentation durations of 50 ms ($r = 0.573$, $p < .001$), 100 ms ($r = 0.566$, $p = .001$), and 500 ms ($r = 0.619$, $p < .001$). Each of these correlations was conserved when controlling for the objective degree of pain intensity in each expression (33 ms: $r = 0.443$, $p < .001$; 50 ms: $r = 0.589$, $p < .001$; 100 ms: $r = 0.559$, $p < .001$; 500 ms: $r = 0.611$, $p < .001$).

Assessing these effects at the level of individual participants weakened these associations (33 ms: avg. Fisher-transformed $r = 0.013$, $SD = 0.117$; 50 ms: avg. Fisher-transformed $r = 0.034$, $SD = 0.145$; 100 ms: avg. Fisher-transformed $r = 0.031$, $SD = 0.149$; 500 ms: avg. Fisher-transformed $r = 0.048$, $SD = 0.168$). That said, the average correlation between individual-level bias and mean criterion ratings was still significantly different from zero within each presentation duration *except* for the 33 ms condition (33 ms: $t(194) = 1.584$, $p = .115$; 50 ms: $t(194) = 3.29$, $p = .001$; 100 ms: $t(194) = 2.94$, $p = .004$; 500 ms: $t(194) = 3.98$, $p < .001$; one-sample *t*-tests versus zero). Taken together, we can conclude that not only is racial bias in pain judgments evident after minimal presentation times, but also that bias evident at least as early as 50 ms is positively correlated with bias in unconstrained judgments.

3.2.4.1. Interactions with perceiver race. Each of the stimulus-level correlations described above generalized across perceiver race, both in terms of the zero-order bivariate correlations (all $rs > 0.223$, all $ps = 0.047$) and the partial correlations controlling for expression intensity (all $rs > 0.241$, all $ps = 0.033$). In both cases, the weakest (though still statistically significant) correlation was observed for White perceivers when seeing faces presented for 33 ms. That said, the strength of this correlation did not differ between Black and White perceivers ($z = 1.12$, $p = .263$).

As for the individual-level correlations, these results were also consistent across perceiver race. While the average Fisher-transformed correlations between judgments at 33 ms and unconstrained judgments did not reach statistical significance for either Black (avg. Fisher-transformed $r = 0.012$, $SD = 0.113$; $t(94) = 1.06$, $p = .293$) or White perceivers (avg. Fisher-transformed $r = 0.014$, $SD = 0.122$; $t(99) = 1.17$, $p = .244$), this relationship was significant at all other presentation durations regardless of perceiver race, with the exception of Black perceivers in the 100 ms presentation condition (avg. Fisher-transformed $r = 0.029$, $SD = 0.147$; $t(99) = 1.94$, $p = .056$; all average Fisher-transformed $rs > 0.028$, all $ps < 0.047$).

3.3. Discussion

The results of Experiment 1B replicate those observed in Experiment 1A. Participants continued to see pain more readily on White versus Black faces after only 33 ms presentation. Here, this bias generalized not only across presentation duration but also, critically, perceiver race: both Black and White participants displayed this bias and if anything, it was magnified in Black participants. As in Experiment 1A, this bias was exacerbated by expression ambiguity—both Black and White participants showed the greatest disparity between pain judgments on Black and White faces for the 50% pain morphs. In sum, these data add confidence to our assertion that racial bias in pain perception occurs rapidly.

4. Experiment 2A: assessing the spontaneity of racial bias in pain perception

Next, we examined whether racial bias in pain perception is dependent on controlled processing, by varying the degree of load participants were under while completing our standard pain perception task. Other work demonstrates that biases in pain care may be primarily driven by automatic (versus controlled) processes (Mathur et al., 2014). Moreover, Experiments 1A-B suggested that perceptual bias in this context is observed even when stimuli are presented rapidly, without sufficient time for inspection or deliberation. As such, we might expect no difference in racial bias in pain perception as a function of load. That said, the pain perception task employed in our previous work is not an entirely implicit task (Drain et al., 2022; Mende-Siedlecki et al., 2021, 2019); it is possible that previous participants were aware that we were measuring their racial bias and as such, they may have been consciously trying to regulate that bias. For this reason, we initially predicted that racial bias in pain perception would be *exacerbated* under conditions of high load.

4.1. Methods

4.1.1. Participants

We recruited 300 Mechanical Turk participants living in the United States (166 men, 133 women, 1 non-binary; M_{age} = 34.50, SD_{age} = 10.16).

We preregistered our procedure, stimuli, sample size, and analysis plan (https://osf.io/g2ez4). We aimed for a sample of 125 White participants, which we determined a priori would afford us sufficient power to detect our smallest predicted effect: a small-to-moderately sized interaction between the effects of target race and cognitive load. As we could not employ a race-based recruitment criterion on MTurk, 75 individuals identifying with a racial or ethnic group besides White were excluded from analyses. We monitored our exclusion criteria during data collection and recruited participants in several waves until we achieved this sample size, though we did not begin analyses until after completion. Besides participant race, we also screened based on VPS/VPN detection as in Experiments 1A-B (Winter et al., 2019). MTurk participants who participated in tasks using similar paradigms or stimuli were also prevented from participating.

We also excluded participants from analysis if they a) displayed non-differentiation of responses (e.g., the same response on 90% of trials or greater), b) indicated in a post-task question assessing their strategy for the cognitive load manipulation that they either wrote down or digitally copied down the patient ID numbers (resulting in 92 exclusions), c) gave an answer to an open-ended post-task question about the study's purpose that approximated the interaction between target race and cognitive load on pain perception thresholds (confirmed if two out of three coders [the first, third, and fourth authors] determined that this rule had been violated; resulting in no exclusions[6]), or d) violated an exclusion rule established by Hackel, Looser, and Van Bavel (2014) for use with this type of paradigm (we averaged participants' responses for morph 1 and morph 11 within each cell of the design, subtracted the former from the latter, and if the resultant was less than 1, the participant was excluded from analysis). In brief, this latter rule effectively excludes participants either giving random responses in the pain task or not consistently distinguishing between painful and neutral expressions. This rule resulted in 8 exclusions (one of which would have also failed the non-differentiation criterion). In total, these criteria left 125 White participants for analyses (64 men, 61 women; M_{age} = 33.54, SD_{age} = 9.47). (Thus, we had 80% power to detect an effect size of $f = 0.126$, for the potential interaction between target race and cognitive load [G*Power; v3.1].)

4.1.2. Stimuli

We selected 12 target heads and 6 expressions from the digitally rendered stimuli in the DPD (Mende-Siedlecki et al., 2020)—specifically, 6 pain expressions that were rated as looking more like pain on average than any other emotion (M = 5.04 out of 7; all comparison emotion $Ms < 3.26$, all comparison $ps < 0.0006$), according to a previous norming of these stimuli. We varied and partially counterbalanced the pairings of target heads, expressions, and load conditions across four versions of the task. Using these targets and expressions, we created 1056 individual stimuli which varied in race (528 Black, 528 White) and pain intensity (each target/expression was created at 11 levels of intensity: from 0% to 100% pain, in 10% increments). Participants were randomly assigned to see all stimuli from one of the four task versions. Within each version, Black and White versions of each target/expression pairing appeared, always within the same load condition. While in Experiments 1A-B we explicitly held structure constant across target race, here, the internal structure features of Black and White faces were allowed to vary in terms of racial prototypicality (Fig. 4). That said, we manually equated these stimuli on width and height, since facial width-to-height ratio may influence pain judgments (Deska & Hugenberg, 2018). Ultimately, each participant rated 264 face stimuli in total.

To confirm that pain was not rendered differently on the Black and White targets, we submitted these stimuli to OpenFace, an open-source deep-learning algorithm (Baltrusaitis, Zadeh, Lim, & Morency, 2018), to automate the process of identifying the activation of five pain-related facial action units—specifically, brow lowering, cheek raising, lid tightening, nose wrinkling, and eye closing (e.g., Kunz et al., 2019). Specifically, we computed how many of these action units, on average, were detected on the stimuli we created for Experiment 2A. The created stimuli did not differ significantly in terms of algorithmic assessment of expressed pain content in the 100% pain expressions (M_{Black} = 4.71, M_{White} = 4.71, $p > .999$).

4.1.3. Design and task

This experiment was conducted on Qualtrics and took the form of a 2 (Target race: Black vs. White) × 2 (Cognitive load: low vs. high) repeated measures design. Participants completed a modified version of an existing paradigm (Mende-Siedlecki et al., 2019), in which they saw 24 male targets (12 Black, 12 White; created using FaceGen Pro) presented on a white background. On average, each face was approximately 2.5 in. by 2.5 in. in size. While we did not request a specific viewing distance in Experiment 2A (which was conducted prior to both

[6] Three participants gave responses that referred to all three of these variables but did not describe an interaction between target race and load. For the sake of comprehensiveness, we assessed whether the results to come were robust to their exclusion, which they were (see Supplementary Materials).

Fig. 4. Example stimuli from Experiment 2A. Participants saw Black (top) and White (bottom) morphs ranging from 0% pain (e.g., neutral; left) to 100% painful (right) facial expressions along 11 equidistant points. The Black and White faces depicted above a) were rendered from the same base head structure (though internal features on Black and White faces vary subtly as a function of racial prototypicality) and b) are making the same painful expression.

Experiment 1A and 1B), assuming a viewing distance of approximately 24 in., each face would have subtended approximately 6.0° both vertically and horizontally. We also did not collect screen resolution information Experiment 2A.

Each target was paired with an ID number to instantiate the cognitive load manipulation. We stated in our instructions, "*To make sure that you're paying attention, we'll be asking you to keep track of each patient's ID number during the experiment. At the beginning of each patients' set of images, you'll see their face and their patient ID number, and you'll have a few seconds to memorize the number. At the end of that set of images, you'll be asked to type in the ID number. Please do not copy down the ID numbers. Please do your best to memorize them without assistance.*"

6 targets appeared in each cell of the design (e.g., 6 Black targets paired with low load ID numbers, etc.). Each target was seen 11 times depicting 11 degrees of pain (morphs increasing from 0% pain to 100% pain). Each target's images were fully contained within their own block. Each block began with a given target's neutral face, a patient ID number, and a reminder that the participant would be asked to provide that ID number at the end of the block. Patient ID numbers were random strings of 2 (low load) or 8 (high load) digits. This manipulation was adapted from prior work (Gilbert & Osborne, 1989; Todorov & Uleman, 2003).

On each trial within a given block, participants viewed a Black or White face depicting some level of pain. The participant's task was to indicate the degree of pain they believed the target face was experiencing (7-point scale ("Definitely not in pain" to "Definitely in pain"). Faces were presented for an unlimited time window; once participants made a response, the next face appeared. Following the presentation and rating of all 11 targets in a given block, the participant was asked to enter the patient ID number of the target they had just been rating. All participants saw all pain levels of all targets' faces exactly one time. Morphs were presented in a randomized order within blocks and blocks were also presented in a randomized order.

Following the task, participants were asked a series of task-specific questions: an open-ended item asking what they thought the study was about, a multiple-choice item assessing their strategy regarding the pain ratings, and critically, a multiple-choice item assessing their strategy regarding the patient ID memorization. Participants who indicated "I copied the numbers down, either on paper or in another window" were excluded from analyses.

4.1.3.1. Individual difference measures. Finally, participants completed a demographics survey, as well as the same exploratory individual difference measures used in Experiment 1. We also collected social evaluations of the physical strength, status, and threat posed by 12 social groups (including Black and White Americans).

4.1.4. Analyses

First, participant ratings were linearly transformed to a scale from 0 to 1 (0 = not in pain, 1 = in pain). Next, ratings from each condition were separately fit with a cumulative normal function to calculate the PSE in each cell of the 2 (target race) × 2 (cognitive load). The PSE represents the point at which a target would be equally likely to be judged as being in pain or not being in pain. This task and analytic approach were ultimately adapted from Looser & Wheatley, 2010 and Hackel et al., 2014. Higher values within a given cell indicate that more pain content was needed on faces from that condition to be judged as being in pain.

We then conducted a 2 (target race) × 2 (cognitive load) repeated measures ANOVA on thresholds for pain perception to test whether a) thresholds for the visual perception of pain expressions varied based on target race and cognitive load, as well b) whether the effect of target race varied as a function of load. We predicted a main effect of target race; participants would see pain less readily on Black (vs. White) faces, overall. In our preregistration, we initially predicted an interaction between target race and load, such that racial bias in pain perception would be more pronounced in the high load (vs. low load) condition. Finally, while we did not preregister an analysis of participants' recollections of targets' ID numbers, one would expect based on prior work using this manipulation that accuracy would be higher in the low load (e.g., 2-digit) condition, compared to the high load (e.g., 8-digit) condition.

4.2. Results

4.2.1. Cognitive load manipulation check

We observed a main effect of load condition on participants' accuracy in remembering targets' patient ID numbers ($F(1,124) = 166.58$; $p < .001$, $\eta_p^2 = 0.57$). Participants were considerably more accurate when recalling 2-digit (e.g., low load) ID numbers ($M = 0.937$, $SD = 0.132$), compared to 8-digit (e.g., high load) ID numbers ($M = 0.584$, $SD = 0.325$), giving us confidence that our load manipulation was successful.

4.2.2. Effects of target race and cognitive load on pain perception

We observed a statistically significant main effect of target race on participants' perceptual thresholds for seeing pain during the task ($F(1,124) = 31.24$; $p < .001$, $\eta_p^2 = 0.20$). Specifically, participants' thresholds for seeing pain on Black faces were more stringent ($M_{Black} = 0.525$, $SD_{Black} = 0.145$; $M_{White} = 0.501$, $SD_{White} = 0.134$) compared to their thresholds for seeing pain on White faces.

That said, neither the interaction between target race and cognitive load ($F(1,124) = 0.24$, $p = .627$, $\eta_p^2 < 0.01$) nor the main effect of cognitive load itself ($F(1,124) < 0.01$, $p = .995$, $\eta_p^2 < 0.01$) was statistically significant. In other words, the effect of target race on participants' perceptual thresholds persisted regardless of whether participants were under low cognitive load ($F(1,124) = 21.36$, $p < .001$, $\eta_p^2 = 0.15$; $M_{Black} = 0.525$, $SD_{Black} = 0.140$; $M_{White} = 0.502$, $SD_{White} = 0.133$) or high cognitive load ($F(1,124) = 19.18$, $p < .01$, $\eta_p^2 = 0.13$; $M_{Black} = 0.526$, $SD_{Black} = 0.156$; $M_{White} = 0.501$, $SD_{White} = 0.141$; Fig. 5).

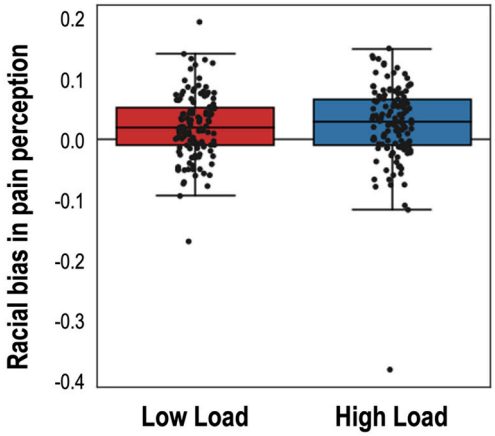

Fig. 5. Racial bias in pain perception, split by cognitive load in Experiment 2A (computer-generated stimuli). Participants rehearsed 2- and 8-digit numbers (low vs. high load) while completing a pain perception task involving computer-generated Black and White faces. Positive values reflect higher perceptual thresholds for detecting pain on Black versus White faces. Indeed, participants saw pain less readily on Black versus White faces independent of cognitive load (red = low load, blue = high load). Box boundaries represent lower and upper quartiles, interior lines within boxes represent the median within each condition, whiskers extend to within 1.5 times the interquartile range of upper and lower quartiles, and dots represent individual bias scores. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Discussion

We observed that cognitive load neither diminished nor enhanced racial bias in pain perception. Instead, participants saw pain less readily on Black (versus White) faces whether they were under high load or comparatively low load. Since the cognitive load manipulation we used interferes with controlled processing and taxes attentional resources, these findings suggest that racial bias in pain perception occurs with relative spontaneity and efficiency.

5. Experiment 2B: confirming the spontaneity of racial bias in pain perception

Given that the findings of Experiment 2A were contrary to our initial prediction that cognitive load would *increase* racial bias in pain perception, we conducted a replication of this protocol in Experiment 2B. In order to test the generalizability of these results, we conducted this replication using a different set of stimuli (e.g., real images of participants posing pain) and in a different sample (e.g., a college undergraduate sample). Based on the results of Experiment 2A, we predicted that target race would influence pain perception regardless of cognitive load.

5.1. Methods

5.1.1. Participants

We recruited 160 participants via the UD PSYC100 subject pool (101 women, 53 men, 6 non-reported; M_{age} = 18.77, SD_{age} = 0.79) who completed this experiment remotely via Qualtrics.

We preregistered our procedure, stimuli, sample size, and analysis plan (https://osf.io/kq65h). We used the same exclusion criteria and data monitoring approach as in Experiment 2A, which resulted in 13 exclusions (6 for task-based criteria; 7 individuals belonging to a racial or ethnic group other than White were also excluded). In addition, three individuals participated more than once; only their initial participation was analyzed. Finally, while we initially sought 125 participants for this experiment (as in Experiment 2A), recruitment continued to the end of the semester, resulting in 13 additional participants. In accordance with our preregistration, only the first 125 White participants were included in analyses. In total, these criteria left 125 White participants for analyses (86 female, 39 male; M_{age} = 18.78, SD_{age} = 0.83). (As such, we once again had 80% power to detect an effect size of f = 0.126, for the potential interaction between target race and cognitive load [G*Power; v3.1].)

5.1.2. Stimuli

We selected 12 Black and 12 White targets (both neutral and painful expressions) from the photographic stimuli portion of the DPD (Mende-Siedlecki et al., 2020). In brief, these stimuli are images of faces (cropped and edited to remove necks, shoulders, and attire) of models posing facial reactions to various painful experiences (e.g., burn, shock, cut, etc.) in a standardized procedure (e.g., multiple images taken at multiple levels of pain).

As in Experiment 2A, we submitted these stimuli to OpenFace (Baltrusaitis et al., 2018) to compute how many pain-related action units, on average, were expressed on both the neutral and painful expressions we chose to include in Experiment 2B. The selected stimuli did not differ significantly in terms of algorithmic assessment of expressed pain content in the 100% pain expressions (M_{Black} = 4.00, M_{White} = 4.00, p > .999) or latent pain content in the neutral expressions (M_{Black} = 0.58, M_{White} = 0.67, p = .797).

Further, norming data collected during the validation of the DPD allowed us to minimize variability across groups on various pain-relevant dimensions. These stimuli did not differ significantly in terms of subjective assessments of pain intensity (M_{Black} = 4.56, M_{White} = 4.69, p = .395), pain specificity (M_{Black} = 1.60, M_{White} = 1.62, p = .919), and believability (M_{Black} = 5.57, M_{White} = 5.47, p = .655) of their painful expressions, the latent pain content of their neutral expressions (M_{Black} = 1.82, M_{White} = 1.97, p = .178), or the strength (M_{Black} = 4.30, M_{White} = 4.07, p = .250), dominance (M_{Black} = 4.03, M_{White} = 3.88, p = .429), trustworthiness (M_{Black} = 3.24, M_{White} = 3.20, p = .819), or perceived status (high status: M_{Black} = 2.72, M_{White} = 3.03, p = .122; low status: M_{Black} = 3.54, M_{White} = 3.24, p = .125) of their neutral images. (For further details, see Supplementary Table 2.)

Using Morpheus PhotoMorpher Pro, we created morphs between each target's neutral and painful expressions at 11 levels of intensity: from 0% to 100% pain, in 10% increments (Fig. 6). We partially counterbalanced pairings of targets to load level and varied the groupings of the targets across four versions of the task. Participants saw all stimuli from one randomly assigned version.

5.1.3. Design and task

This experiment took the form of a 2 (Target race: Black vs. White) × 2 (Cognitive load: low versus high load) repeated measures design and was identical in structure to the task in Experiment 2A. 6 targets once again appeared in each cell of the design (e.g., 6 White targets paired with high load ID numbers, etc.). The faces of these targets were presented against a white background. We once again requested (as in Experiments 1A-B) that participants maintain a viewing distance of 24 in. from their screens. On average, each face was approximately 3 in. tall and 2.25 in. wide. As such, each face subtended approximately 7.2° vertically and 5.4° horizontally. Within the participants included in analyses, the modal screen resolution was 1440 × 900, with an average of 1457.87 × 891.34.

5.1.4. Analyses

While our data preprocessing and analyses steps were identical to Experiment 2A, our predictions shifted slightly based on the previous results we observed. Specifically, we predicted that while target race would exert its typical influence on thresholds for pain perception (e.g., participants would have more stringent thresholds for seeing pain on Black versus White faces), target race would not interact with cognitive load.

Fig. 6. Example stimuli from Experiments 2B and 2-C. Participants saw Black (top) and White (bottom) morphs ranging from 0% pain (e.g., neutral; left) to 100% painful (right) facial expressions along 11 equidistant points. Individuals depicted here granted full permission for their likenesses to appear in this article.

5.2. Results

5.2.1. Cognitive load manipulation check

As in Experiment 2A, we observed a main effect of load condition on participants' accuracy in remembering the targets' patient ID numbers ($F(1,124) = 562.67$; $p < .001$, $\eta_p^2 = 0.81$). Participants were more accurate when recalling 2-digit (e.g., low load) ID numbers ($M = 0.848$, $SD = 0.214$), compared to 8-digit (e.g., high load) ID numbers ($M = 0.318$, $SD = 0.246$), indicating that our load manipulation was again successful.

5.2.2. Effects of target race and cognitive load on pain perception

Replicating our findings in Experiment 2A, we observed a statistically significant main effect of target race on participants' perceptual thresholds for seeing pain during the task ($F(1,124) = 18.94$; $p < .001$, $\eta_p^2 = 0.13$). Participants' thresholds for seeing pain on Black faces were more stringent compared to their thresholds for White faces ($M_{Black} = 0.478$, $SD_{Black} = 0.146$; $M_{White} = 0.456$, $SD_{White} = 0.124$).

Also replicating Experiment 2A, neither the interaction between target race and cognitive load ($F(1,124) = 0.65$, $p = .421$, $\eta_p^2 < 0.01$) nor the main effect of load itself ($F(1,124) < 0.01$, $p = .944$, $\eta_p^2 < 0.01$) was statistically significant. In other words, participants' perceptual thresholds were more stringent for pain on Black (versus White) faces regardless of whether participants were under low cognitive load ($F(1,124) = 13.49$; $p < .001$, $\eta_p^2 = 0.10$; $M_{Black} = 0.480$, $SD_{Black} = 0.150$; $M_{White} = 0.455$, $SD_{White} = 0.134$) or high cognitive load ($F(1,124) = 6.33$; $p < .001$, $\eta_p^2 = 0.05$; $M_{Black} = 0.476$, $SD_{Black} = 0.150$; $M_{White} = 0.458$, $SD_{White} = 0.133$; Fig. 7).

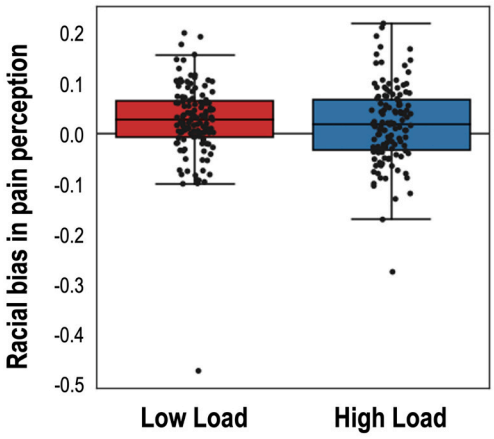

Fig. 7. Racial bias in pain perception, split by cognitive load in Experiment 2B (photographic stimuli). Participants once again saw pain less readily on Black versus White faces independent of cognitive load (red = low load, blue = high load). Box boundaries represent lower and upper quartiles, interior lines within boxes represent the median within each condition, whiskers extend to within 1.5 times the interquartile range of upper and lower quartiles, and dots represent individual bias scores. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.3. Discussion

This replication adds strength to our conclusions from Experiment 2A. Once again, cognitive load did not affect racial bias in pain perception; participants saw pain less readily on Black (versus White) faces whether they were under high load or comparatively low load. At the same time, it is also notable that load did not *enhance* this perceptual bias either. Furthermore, Experiment 2B demonstrates the generalizability of these results by employing a novel set of stimuli (e.g., real images of humans posing pain instead of the computer-generated faces used in Experiment 2A) and recruiting a different demographic sample (e.g., college undergraduates).

6. Experiment 2C: strong evidence for the spontaneity of racial bias in pain perception

While Experiments 2A-B suggest that racial bias in pain perception operates with relative automaticity, several alternative explanations potentially remain. First, while we took differences in accuracy between the load conditions as a sign that it was harder to remember the 8-digit numbers and, as a result, that participants were exerting greater effort in the high load condition, we did not explicitly measure difficulty or effort. At minimum, it would be useful to screen out participants who report exerting less effort in the high load condition or who did not find the high load condition to be more difficult than the low load condition. Moreover, given that these tasks were self-paced, participants may have spent more time attending to faces in one load condition versus the other. As such, in Experiment 2C, we time-locked the presentation duration of each face, as well as the corresponding response windows. Lastly, while Experiments 2A-B did not reveal significant differences in perceptual bias between high and low load conditions, this comparison rests on inferences made from a null result. Conducting Experiment 2C gave us the opportunity to preregister a secondary Bayesian analysis comparing differences in racial bias in pain perception in the high and low load conditions, which can offer greater evidence in support of the null hypothesis. Taken together, Experiment 2C offers an even stronger test of our hypothesis that racial bias in pain perception is spontaneous.

6.1. Methods

6.1.1. Participants

We recruited 255 White participants via Prolific (156 women, 84 men, 10 non-binary, 5 not reported; $M_{age} = 29.54$, $SD_{age} = 10.44$).

We preregistered our procedure, stimuli, sample size, and analysis plan (https://osf.io/v938j). We expanded on the exclusion criteria and data monitoring approach as in Experiments 2A-B. In addition to the procedures implemented there, we also asked several additional questions immediately post-task regarding participants' experience of task difficulty and effort (e.g., whether or not they found it more difficult to remember the patient ID numbers in certain conditions, whether or not they tried harder in certain conditions). We stated in our pre-registration that we would exclude participants who indicated that a) it was harder to remember the shorter ID numbers, compared to the longer ID numbers or b) it was equally hard to remember both types of numbers, or who indicated that c) they tried harder to remember the shorter ID numbers than the longer ID numbers or d) they tried to remember the shorter ID numbers, but gave up at trying to remember the longer ID numbers.

This approach resulted in 130 exclusions (40 whose responses violated the exclusion criteria specific to the pain perception task; 16 who did not respond to the minimum number of patient ID items; 60 based on post-task responses regarding task difficulty, effort, attention, disruptions, or who said that they copied the ID numbers manually). In addition, 3 individuals participated more than once (only their initial participation was analyzed), while 2 individuals did not complete the post-task and therefore could not be included in our analyses. Finally, while we initially sought 125 participants for this experiment (as in Experiment 2A), 9 additional individuals participated. In accordance with our preregistration, only the first 125 White participants were included in analyses.

In total, these criteria left 125 White participants for analyses (97 women, 21 men, 7 non-reported; $M_{age} = 26.31$, $SD_{age} = 8.69$). (As such, we once again had 80% power to detect an effect size of $f = 0.126$, for the potential interaction between target race and cognitive load [G*Power; v3.1].)

6.1.2. Stimuli

We selected both neutral and painful expressions from 12 Black and 12 White targets from the DPD (Mende-Siedlecki et al., 2020). The selected stimuli did not differ significantly in terms of algorithmic assessment of expressed pain content in the 100% pain expressions ($M_{Black} = 3.92$, $M_{White} = 3.50$, $p = .406$) or latent pain content in the neutral expressions ($M_{Black} = 0.50$, $M_{White} = 0.58$, $p = .784$). Moreover, these stimuli did not differ significantly in terms of subjective assessments of pain intensity ($M_{Black} = 4.48$, $M_{White} = 4.68$, $p = .306$), pain specificity ($M_{Black} = 1.34$, $M_{White} = 1.69$, $p = .211$), and believability ($M_{Black} = 5.58$, $M_{White} = 5.40$, $p = .459$) of their painful expressions, the latent pain content of their neutral expressions ($M_{Black} = 1.80$, $M_{White} = 1.95$, $p = .187$), or the strength ($M_{Black} = 4.18$, $M_{White} = 4.05$, $p = .515$), dominance ($M_{Black} = 3.87$, $M_{White} = 3.86$, $p = .954$), trustworthiness ($M_{Black} = 3.39$, $M_{White} = 3.15$, $p = .180$), or perceived status (high status: $M_{Black} = 2.78$, $M_{White} = 2.95$, $p = .350$; low status: $M_{Black} = 3.45$, $M_{White} = 3.29$, $p = .407$) of their neutral images. (For further details, see Supplementary Table 3.) Morph between neutral and painful expressions were once again created using Morpheus PhotoMorpher Pro. Counterbalancing and randomization was identical to Experiment 2B.

6.1.3. Design and task

This experiment took the form of a 2 (Target race: Black vs. White) × 2 (Cognitive load: low versus high load) repeated measures design and was identical in structure to the task in Experiments 2A-B. 6 targets once again appeared in each cell of the design (e.g., 6 White targets paired with high load ID numbers, etc.). The faces of these targets were presented against a gray background and participants were again asked to maintain a viewing distance of 24 in. from their screens. On average, each face was approximately 3 in. tall and 2.25 in. wide. As such, each face subtended approximately 7.2° vertically and 5.4° horizontally. Within the participants included in analyses, the modal screen resolution was 1440 × 900, with an average of 1512.47 × 893.21.

In addition to the new screening questions described above (which were asked immediately following completion of the main task), we note two other key changes. First, in order to equate both face presentation and decision duration across cognitive load and across target race, we presented faces for a fixed duration of 4 s. Participants had only this fixed window during which to respond and rate how painful each face's expression was, followed by an intertrial interval of 1 s. Further, at the end of each block of morphs, participants had a fixed 10 s window in which to type in that particular target's patient ID number. To accommodate this specificity of timing, we conducted the main task through PsychoPy v2021.1.2 (Peirce, 2007), though the post-task surveys were still completed via Qualtrics.

6.1.4. Analyses

As in Experiment 2B, we predicted that while target race would exert its typical influence on thresholds for pain perception (e.g., participants would have more stringent thresholds for seeing pain on Black versus White faces), target race would not interact with cognitive load.

To provide stronger evidence for the null hypothesis that racial bias in pain perception is not moderated by cognitive load, we conducted a separate Bayesian analysis of these data. Specifically, we ran a Bayesian one-way repeated measures ANOVA in SPSS comparing racial bias in pain perception within the low load condition versus the high load condition. This approach derives a Bayes factor from the Bayesian information criterion, in order to quantify support for the null hypothesis (that bias does not differ as a function of load) versus the experimental hypothesis (that bias *does* differ as a function of load). We interpreted the strength of the resulting Bayes factor using commonly accepted cutoff values (Jarosz & Wiley, 2014).

6.2. Results

6.2.1. Cognitive load manipulation check

As in Experiment 2A and 2B, we observed a main effect of load condition on participants' accuracy in remembering the targets' patient ID numbers ($F(1,124) = 742.28$; $p < .001$, $\eta_p^2 = 0.86$). Participants were more accurate when recalling 2-digit (e.g., low load) ID numbers ($M = 0.877$, $SD = 0.094$), compared to 8-digit (e.g., high load) ID numbers ($M = 0.288$, $SD = 0.245$), indicating that our load manipulation was again successful.

6.2.2. Effects of target race and cognitive load on pain perception

Replicating our findings in Experiment 2A and 2B, we observed a statistically significant main effect of target race on participants' perceptual thresholds for seeing pain during the task ($F(1,124) = 27.14$; $p < .001$, $\eta_p^2 = 0.18$). Participants' thresholds for seeing pain on Black faces were more stringent compared to their thresholds for White faces ($M_{Black} = 0.534$, $SD_{Black} = 0.172$; $M_{White} = 0.503$, $SD_{White} = 0.171$).

Also replicating Experiment 2A and 2B, neither the interaction between target race and cognitive load ($F(1,124) = 0.131$, $p = .718$, $\eta_p^2 < 0.01$) nor the main effect of load itself ($F(1,124) = 0.296$, $p = .588$, $\eta_p^2 < 0.01$) was statistically significant. Once again, participants' perceptual thresholds were more stringent for pain on Black (versus White) faces regardless of whether participants were under low cognitive load ($F(1,124) = 11.40$; $p < .001$, $\eta_p^2 = 0.084$; $M_{Black} = 0.534$, $SD_{Black} = 0.182$; $M_{White} = 0.499$, $SD_{White} = 0.175$) or high cognitive load ($F(1,124) = 8.224$; $p = .005$, $\eta_p^2 = 0.062$; $M_{Black} = 0.535$, $SD_{Black} = 0.176$; $M_{White} = 0.507$, $SD_{White} = 0.193$; Fig. 8).

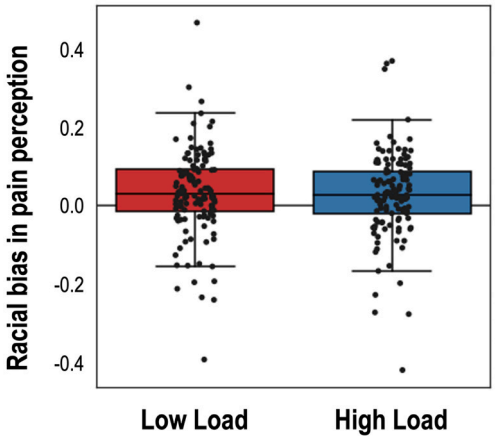

Fig. 8. Racial bias in pain perception, split by cognitive load in Experiment 2C (photographic stimuli). Participants continued to see pain less readily on Black versus White faces independent of this load manipulation (red = low load, blue = high load). Box boundaries represent lower and upper quartiles, interior lines within boxes represent the median within each condition, whiskers extend to within 1.5 times the interquartile range of upper and lower quartiles, and dots represent individual bias scores. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6.2.3. *Bayesian analyses of the effect of load on racial bias in pain perception*

When comparing racial bias in pain perception in the low load condition versus the high load condition using a Bayesian one-way repeated measures ANOVA, we observed a BF_{01} of 0.069. Given the typical cutoff values associated with Bayes factors (Jarosz & Wiley, 2014), this represents strong evidence in favor of the null hypothesis that racial bias in pain perception does *not* vary as a function of cognitive load.

We also revisited the datasets from Experiments 2A and 2B and performed similar analyses. In both cases, the Bayes factors observed for this comparison ($BF_{01\_Exp2A} = 0.081$, $BF_{01\_Exp2B} = 0.089$) supported similar conclusions. Finally, we conducted one additional analysis collapsing across all three experiments, which once again yielded strong evidence in favor of the null hypothesis ($BF_{01} = 0.043$).

6.3. *Discussion*

Using a more conservative screening approach, more tightly controlled presentation conditions, and a more appropriate statistical analysis, we continued to observe no evidence of a difference in racial bias in pain perception as a function of cognitive load. Participants continued to see pain less readily on Black faces regardless of high or low load. Furthermore, Bayesian analysis (both within this study and across the previous two) indicated strong evidence for the null hypothesis that perceptual bias did not differ across these two conditions. In sum, Experiments 2A-C demonstrate that target race influences pain perception independent of cognitive load, suggesting that effortful, deliberative processing is not required for this bias to be observed. In other words, racial bias in pain perception is a relatively automatic phenomenon.

7. General discussion

Previous work demonstrates that biases in the basic visual perception of painful expressions may be linked to pervasive racial disparities in pain care. Here, we examined the speed and spontaneity of this racial bias in pain perception. Across two experiments, we observed that this perceptual bias operates rapidly—participants were more likely to judge White (versus Black) faces as being in pain after only 33 ms of presentation. Bias under minimal presentation conditions also predicted bias at unconstrained presentation.[7] Furthermore, these findings were generalizable across perceiver race, as both Black and White perceivers showed evidence of this perceptual bias. Across three additional experiments, we also observed that this perceptual bias operates with relative automaticity, rather than depending on effortful, deliberative processing.

The rapid emergence of bias in pain perception is line with other work demonstrating strong race-based influences on emotion perception and categorization at minimal presentation times (e.g., Craig et al., 2012). For example, Craig and colleagues observed that a happy categorization advantage for White faces and an angry categorization advantage for Black faces were both robust when faces were presented for 200 ms, but that these effects were not observed within unlimited presentations—at least within computer-generated stimuli. These authors interpret this finding through the lens of work on visual attention (Lavie & De Fockert, 2003) suggesting that when sensory input is minimal, the influence of task-irrelevant information (in this case, target race) is enhanced. However, the present data differ from these findings in that here, bias in pain judgments was *not* weakened for unlimited presentation durations, nor did it vary as a function of presentation duration at all. One possibility is that the effect of target race on pain perception is simply stronger than effects that have been tested within the context of other emotions. Indeed, when compared against similar gaps in recognizing anger, happiness, and fear, racial bias in pain perception is particularly robust (Mende-Siedlecki et al., 2021).

Notably, the effect of target race on pain judgments in Experiments 1A-B was strongly moderated by expression intensity. Racial bias was largest for the most ambiguous stimuli—the 50% morphs—and scaled down as targets' facial expressions became more obviously neutral or painful. This pattern was also observed in Supplementary Experiment 1. More broadly, this data dovetails with a rich literature concerning top-down influences on social perception. Contextual information (including cues to race, status, or group membership) exerts its strongest influence on perceptually ambiguous stimuli (Freeman & Johnson, 2016; Pauker, Rule, & Ambady, 2010; Xiao et al., 2016). In other words, as bottom-up cues become more uncertain, it becomes more likely that these gaps are filled in from the top down. This tendency has clear import in clinical settings, where research demonstrates that perceiver/patient correspondence is weakest for submaximal pain (Prkachin, Berzins, & Mercer, 1994).

Experiment 1B also offered a direct comparison of the effect of target race on pain judgments between Black and White perceivers, the first such comparison in our own work. Notably, both Black and White perceivers showed a tendency to rate painful expressions on White faces as looking more like pain than the same expressions of pain rendered on Black faces. This finding is consistent with previous work (e.g., Trawalter et al., 2012; where both Black and White participants expected Black [versus White] targets to experience less pain across various injuries), as well as meta-analyses across our own work ($N = 40$ experiments) on biases in the visual perception of pain (Lin et al., 2021, where Black participants show a racial bias in pain perception that is similar in magnitude and direction to Asian, Hispanic White, and non-Hispanic White participants). Indeed, if anything, Black participants in Experiment 1B demonstrated a *greater* degree of racial bias in pain judgments compared to White participants, echoing the findings of Deska et al. (2020) in the domain of social pain.

On the one hand, it's possible that Black perceivers demonstrated this tendency because of *positive* stereotypes about their in-group regarding resilience and toughness (as suggested by Deska et al.,

[7] In Experiment 1B, the relationship between racial bias in pain judgments at 33 ms presentation and bias at unconstrained presentation was not statistically significant when assessed at the subject level. That said, this relationship was observed at 33 ms a) when assessed at the stimulus level in Experiment 1B and b) when assessed in either manner in Experiment 1A.

2020). From this perspective, the same bias may manifest from different sources in Black and White perceivers. On the other hand, this pattern of data is also broadly consistent with other work in which minoritized individuals demonstrate biases in judgment and perception that are consistent with those held by out-group majority members. For example, many Black participants demonstrate similar implicit racial biases as White participants (e.g., Dasgupta, 2004; Livingstone, 2002; Richeson, Trawalter, & Shelton, 2005), potentially stemming from internalized negative stereotypes about their racial group and knowledge of its place in the social hierarchy (Ashburn, Knowles, & Monteith, 2003). From *this* perspective, this bias may share a common source in Black and White perceivers. Ultimately, while more work is needed to confirm whether similar mechanisms support this perceptual bias in Black and White perceivers, the findings of Experiment 1B are consistent with the suggestion that racial bias in pain perception is not merely a function of in-group preference (e.g., Exp. 7 in Mende-Siedlecki et al., 2019; analyses of Black, Asian, and non-Hispanic White participants in Lin et al., 2021).

We also demonstrated that this perceptual bias is maintained when attentional resources and effortful processing are disrupted via cognitive load. Our load manipulation has been extensively used in prior work to disrupt effortful processing (Gilbert & Osborne, 1989; Todorov & Uleman, 2003) and manipulation check data suggested that the difference in load was strong. Of course, clinicians working in environments like emergency departments are under considerably more extreme conditions of load, stress, and fatigue, with demonstrable consequences for treatment and decision-making (; Burgess, Nelson, et al., 2014; Laxmisan et al., 2007; Westbrook, Raban, Walter, & Douglas, 2018). Some work also posits that load in these contexts magnifies racial disparities in healthcare and further, that minoritized patients are disproportionately likely to be served by providers working in high-load care settings (Burgess, 2010). Ultimately, while Experiments 2A-C demonstrate the spontaneity of racial bias in pain perception across both computer-generated and photographic stimuli, future work could test whether even more taxing and stressful contexts exacerbate this bias.

7.1. Limitations & implications

Across this work, we preregistered our protocols and predictions, employed established paradigms and manipulations, and attended to the generalizability with regards to stimuli and samples. While these strengths add confidence to our conclusions, we note several limitations.

First, while our first three experiments (as well as Supplementary Experiment 1) offer consistent evidence of the speed and spontaneity of racial bias in pain perception, questions regarding the underlying mechanisms may best be answered with the incorporation of neuroimaging approaches. For example, it would be advantageous to confirm the temporal dynamics of this perceptual bias by measuring event-related potentials (ERPs) evoked by Black and White faces displaying pain. The finding that this bias operates automatically (rather than via controlled, effortful processing) is broadly consistent with other work demonstrating that perceivers' differential sensitivity to pain on Black (versus White) faces is supported by disruptions in configural face processing (Mende-Siedlecki et al., 2019).[8] That said, do these disruptions exert their influence at the lowest levels of face perception (Hancock & Rhodes, 2008; Rhodes, Hayward, & Winkler, 2006) or at the level of higher-order emotion perception (which is also supported by configural processing; Bombari et al., 2013; Calder, Young, Keane, & Dean, 2000)? Here, functional magnetic resonance imaging (fMRI) could elucidate the nature of these automatic divergences in perception, as well as their downstream consequences for empathy for pain, whose neural signatures have been well-established (Lamm, Decety, & Singer, 2011).

Another important limitation concerns our choice of stimuli. The stimuli in these five experiments were either computer-generated faces (Exps. 1A-B, Exp. 2A) or photographs of real individuals posing painful expressions (Exps. 2B—C). In other words, in neither case were participants evaluating images of real people genuinely experiencing physical pain. Since real and computer-generated face stimuli may be processed differently (e.g., Gaither, Chen, Pauker, & Sommers, 2019; MacDorman, Green, Ho, & Koch, 2009), it stands to reason that racial bias in pain perception might therefore manifest differently as a function of stimulus type. However, meta-analysis suggests that the magnitude of the perceptual bias under study in the present work does not vary based on the stimuli used (Lin et al., 2021). Across forty studies, participants showed similar tendencies to underperceive pain on Black (versus White) faces regardless of whether the stimuli employed were real or computer-generated. As to the question of pain authenticity, perceivers often struggle to distinguish between genuine and posed expressions of pain (Littlewort, Bartlett, & Lee, 2009; Poole & Craig, 1992), suggesting (at least indirectly) that we would observe similar results within genuine (rather than posed) pain expressions. At the same time, recent work demonstrates that both White and Black perceivers are more accurate at distinguishing posed from genuine expressions on White (versus Black) faces (Lloyd, Lloyd, McConnell, & Hugenberg, 2021), implying that sensitivity to pain authenticity in particular may have downstream consequences for racial disparities in pain care. Ultimately, using stimuli that depict real patients or participants in genuine pain—potentially making dynamic, rather than static displays—would add considerably to this growing literature.

Moreover, as in several previous investigations, we used only male targets across this research. Some of our work has observed that this perceptual bias may be larger within male (versus female) targets (Goharzad et al., 2022), potentially because Black men are seen as more racially prototypic than Black women (Goff, Thomas, & Jackson, 2008; Purdie-Vaughns & Eibach, 2008). (Notably, in those studies, Black female targets' pain *was* perceived less readily than the pain of White *men* and importantly, they received the lowest pain reliever prescriptions of all target types; Goharzad et al., 2022). To better understand the speed and spontaneity of racial bias in pain perception, we ultimately chose to study it here under the conditions in which this perceptual bias has been observed most consistently. Similarly, we once again presented faces independent of any other information that could be diagnostic regarding pain experience. Of course, in clinical contexts, pain is not assessed by facial expression alone. Rather, self-reported pain severity is often

[8] One implication here is that configural face processing is, itself, automatic. We recognize that there is likely some room for debate surrounding this proposition. On the one hand, there is evidence to suggest that face-selective responses (for example, the N170 component, which has been linked to configural face processing [e.g., Eimer, Gosling, Nicholas, & Kiss, 2011]) are intact for faces presented outside of conscious awareness (e.g., Pesciarelli, Sarlo, & Leo, 2011; Suzuki & Noguchi, 2013), specifically sensitive to configural information on faces outside of awareness (e.g., Lyyra et al., 2014), and automatic and unmodulated by load (Schindler, Tirloni, Bruchmann, & Straube, 2021; see also Hine et al., 2011). On the other hand, other work casts at least some doubt on the automaticity of configural face processing (e.g., Jackson & Raymond, 2006; Shafto & Pitts, 2015).

considered the "gold standard" clinical measure of a patient's pain experience (Karcioglu, Topacoglu, Dikme, & Dikme, 2018; Pasero & McCaffery, 2010). Future work should address whether the presence of self-reported pain experience influences the magnitude of racial bias in pain perception.

Finally, given our reliance on online data collection (even within our student samples), we were unable to definitively control viewing distance, and therefore visual angle associated with our face stimuli. However, we gave participants explicit directions regarding viewing distance in all but one of these experiments. Given that we screened participants based on their self-reported violations of directions regarding focus, viewing conditions, and interruptions, it is likely that compliance regarding viewing distance was high. It is also worth noting that the *approximate* visual angle associated with our stimuli is consistent with the size of faces during naturalistic viewing in everyday life (Oruc, Shafai, Murthy, Lages, & Ton, 2019). Moreover, while some data suggests that viewing distance is inversely related to perceived expression intensity (Gerhardsson, Högman, & Fischer, 2015), this effect would be constant across the critical within-subjects comparisons throughout these experiments (e.g., Black vs. White faces, high vs. low load). On a related note, other work suggests that differences in face processing as a function of target race (specifically, the own race bias) may be invariant across viewing distances (Lampinen, Roush, Erickson, Moore, & Race, 2015). Ultimately, while a lack of control over viewing distance likely added noise to our observations, it is unlikely to explain any of the within-subjects differences we observed.

In sum, we demonstrated that racial bias in pain perception occurs automatically, based on minimal visual input, and above and beyond other diagnostic information. Addressing perceptual contributions to pain disparities in clinical contexts will require accounting for their automatic nature and understanding when this bias is most likely to emerge. For example, these data suggest that education- or information-based approaches that target consciously-held and applied beliefs may not be sufficient. Instead, these data speak to how fundamentally deep the roots of racism are: racial bias in pain perception is evident after 33 milliseconds of exposure to a face and its manifestation is effortless. As such, it seems clear that *any* individual-focused intervention alone is unlikely to close gaps in pain care. Rather, these gaps must be addressed with long-term structural and systemic change.

Open practices

All data, stimuli, and supporting materials (e.g., stimulus norming details) associated with all experiments in this manuscript are available on the Open Science Framework (https://osf.io/ht2u8). All experiments were preregistered on OSF as well (Experiment 1A: https://osf.io/8jczn; Experiment 1B: https://osf.io/xfyjv; Experiment 2A: https://osf.io/g2ez4; Experiment 2B: https://osf.io/kq65h; Experiment 2C: https://osf.io/v938j).

Author note

This research was supported by grants from the National Science Foundation (BCS-1918325 to PMS) and the National Institute of General Medical Sciences (5P20GM103653-08; Junior Investigator: PMS). The authors wish to acknowledge the hard work of the undergraduate research assistants of the Mende-Siedlecki Lab at the University of Delaware, as well as Leor Hackel, Jim Hoffman, and Tim Vickery, for their helpful advice and support. Portions of these findings were presented at the annual meetings of the Society for Personality and Social Psychology and the American Psychological Society. We have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2022.104315.

References

Anderson, K. O., Richman, S. P., Hurley, J., Palos, G., Valero, V., Mendoza, T. R., … Cleeland, C. S. (2002). Cancer pain management among underserved minority outpatients: Perceived needs and barriers to optimal control. *Cancer, 94*(8), 2295–2304.

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition, 21*(1), 61–87.

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L. P. (2018, May). Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)* (pp. 59–66). IEEE.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6*(2), 269.

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology, 66*(12), 2426–2442.

Burgess, D. J. (2010). Are providers more likely to contribute to healthcare disparities under high levels of cognitive load? How features of the healthcare setting may lead to biases in medical decision making. *Medical Decision Making, 30*(2), 246–257.

Burgess, D. J., Nelson, D. B., Gravely, A. A., Bair, M. J., Kerns, R. D., Higgins, D. M., … Partin, M. R. (2014). Racial differences in prescription of opioid analgesics for chronic noncancer pain in a national sample of veterans. *The Journal of Pain, 15*(4), 447–455.

Burgess, D. J., Phelan, S., Workman, M., Hagel, E., Nelson, D. B., Fu, S. S., … van Ryn, M. (2014). The effect of cognitive load and patient race on physicians' decisions to prescribe opioids for chronic low back pain: A randomized trial. *Pain Medicine, 15*(6), 965–974.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527.

Cloutier, J., Li, T., & Correll, J. (2014). The impact of childhood experience on amygdala response to perceptually familiar black and white faces. *Journal of Cognitive Neuroscience, 26*(9), 1992–2004.

Craig, B. M., Mallan, K. M., & Lipp, O. V. (2012). The effect of poser race on the happy categorization advantage depends on stimulus type, set size, and presentation duration. *Emotion, 12*(6), 1303.

Craig, K. D., Versloot, J., Goubert, L., Vervoort, T., & Crombez, G. (2010). Perceiving pain in others: Automatic and controlled mechanisms. *The Journal of Pain, 11*(2), 101–108.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and White faces. *Psychological Science, 15*(12), 806–813.

Dasgupta, N. (2004). Implicit ingroup favoritism, outgroup favoritism, and their behavioral manifestations. *Social justice research, 17*(2), 143–169.

Deska, J. C., & Hugenberg, K. (2018). Targets' facial width-to-height ratio biases pain judgments. *Journal of Experimental Social Psychology, 74*, 56–64.

Deska, J. C., Kunstman, J., Lloyd, E. P., Almaraz, S. M., Bernstein, M. J., Gonzales, J. P., & Hugenberg, K. (2020). Race-based biases in judgments of social pain. *Journal of Experimental Social Psychology, 88*, Article 103964.

Drain, A., Goharzad, A., Qu-Lee, J., Lin, J., & Mende-Siedlecki, P. (2022). Skin deep: How racially prototypic features exacerbate bias in pain perception. Preprint available on PsyArxiv https://psyarxiv.com/6gu8j/ (in prep).

Eimer, M., Gosling, A., Nicholas, S., & Kiss, M. (2011). The N170 component and its links to configural face processing: a rapid neural adaptation study. *Brain research, 1376*, 76–87.

Feagin, J., & Bennefield, Z. (2014). Systemic racism and US health care. *Social Science & Medicine, 103*, 7–14.

Fiscella, K., & Sanders, M. R. (2016). Racial and ethnic disparities in the quality of health care. *Annual Review of Public Health, 37*, 375–394.

Forbes, C. E., Cox, C. L., Schmader, T., & Ryan, L. (2012). Negative stereotype activation alters interaction between neural correlates of arousal, inhibition and cognitive control. *Social Cognitive and Affective Neuroscience, 7*(7), 771–781.

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

Freeman, J. B., Penner, A. M., Saperstein, A., Scheutz, M., & Ambady, N. (2011). Looking the part: Social status cues shape race perception. *PLoS One, 6*(9), Article e25107.

Gaither, S. E., Chen, J. M., Pauker, K., & Sommers, S. R. (2019). At face value: psychological outcomes differ for real vs. computer-generated multiracial faces. *The Journal of social psychology, 159*(5), 592–610.

Gee, G. C., & Ford, C. L. (2011). Structural racism and health inequities: Old issues, new directions. *Du Bois Review: Social Science Research on Race, 8*(1), 115.

Gerhardsson, A., Högman, L., & Fischer, H. (2015). Viewing distance matter to perceived intensity of facial expressions. *Frontiers in Psychology, 6*, 944.

Gilbert, D. T., & Osborne, R. E. (1989). Thinking backward: Some curable and incurable consequences of cognitive busyness. *Journal of Personality and Social Psychology, 57*(6), 940.

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology, 94*(2), 292.

Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). "Ain't I a woman?": Towards an intersectional approach to person perception and group-based harms. *Sex Roles, 59*(5–6), 392–403.

Goharzad, A., Drain, A., Qu-Lee, J., Lin, J., & Mende-Siedlecki, P. (2022). Invisible identities, invisible pain: How race and gender prototypicality inform pain perception & treatment. Preprint available on PsyArxiv https://psyarxiv.com/6gu8j/ (in prep).

González-Roldan, A. M., Martínez-Jauand, M., Muñoz-García, M. A., Sitges, C., Cifre, I., & Montoya, P. (2011). Temporal dissociation in the brain processing of pain and anger faces with different intensities of emotional expression. *PAIN®, 152*(4), 853–859.

Green, C. R., Anderson, K. O., Baker, T. A., Campbell, L. C., Decker, S., Fillingim, R. B., … Vallerand, A. H. (2003). The unequal burden of pain: Confronting racial and ethnic disparities in pain. *Pain Medicine, 4*(3), 277–294.

Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology, 52*, 15–23.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*(1), 45–56.

Hebart, M. (2009). IMScramble. Retrieved from http://martin-hebart.de/code/imscramble.m.

Hine, K., Nouchi, R., & Itoh, Y. (2011). Influence of subjective difficulty on the degree of configural and featural processing in face recognition 1. *Japanese Psychological Research, 53*(3), 246–257.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences, 113*(16), 4296–4301.

Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences, 13*(12), 524–531.

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.

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience, 5*(1), 21–36.

Jackson, M. C., & Raymond, J. E. (2006). The role of attention and familiarity in face identification. *Perception & Psychophysics, 68*(4), 543–557.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving, 7*(1), 2.

Karcioglu, O., Topacoglu, H., Dikme, O., & Dikme, O. (2018). A systematic review of the pain scales in adults: Which to use? *The American Journal of Emergency Medicine, 36*(4), 707–714.

Katz, N. (2002). The impact of pain management on quality of life. *Journal of Pain and Symptom Management, 24*(1), S38–S47.

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods, 8*(4), 614–629.

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology, 109*(5), 901.

Kubota, J. T., & Ito, T. (2017). Rapid race perception despite individuation and accuracy goals. *Social Neuroscience, 12*(4), 468–478.

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

Kunz, M., Meixner, D., & Lautenbacher, S. (2019). Facial muscle movements encoding pain—A systematic review. *Pain, 160*(3), 535–549.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage, 54*(3), 2492–2502.

Lampinen, J. M., Roush, A., Erickson, W. B., Moore, K. N., & Race, B. (2015). The effects of simulated distance on recognition of same race and other race faces. *Visual Cognition, 23*(6), 678–698.

Lavie, N., & De Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics, 65*(2), 202–212.

Laxmisan, A., Hakimzada, F., Sayan, O. R., Green, R. A., Zhang, J., & Patel, V. L. (2007). The multitasking clinician: Decision-making and cognitive demand during and after team handoffs in emergency care. *International Journal of Medical Informatics, 76*(11–12), 801–811.

Lee, P., Le Saux, M., Siegel, R., Goyal, M., Chen, C., Ma, Y., & Meltzer, A. C. (2019). Racial and ethnic disparities in the management of acute pain in US emergency departments: meta-analysis and systematic review. *The American Journal of Emergency Medicine, 37*(9), 1770–1777.

Lin, J., Drain, A., Goharzad, A., & Mende-Siedlecki, P. (2021). What factors fuel racial bias in pain perception and treatment?: A meta-analysis across 40 experimental studies. Preprint available on PsyArxiv https://psyarxiv.com/nky37/.

Littlewort, G. C., Bartlett, M. S., & Lee, K. (2009). Automatic coding of facial expressions displayed during posed and genuine pain. *Image and Vision Computing, 27*(12), 1797–1803.

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology, 38*(4), 405–413.

Lloyd, E. P., Lloyd, A. R., McConnell, A. R., & Hugenberg, K. (2021). Race deficits in pain detection: Medical providers and laypeople fail to accurately perceive pain authenticity among black people. *Social Psychological and Personality Science*. https://doi.org/10.1177/19485506211045887

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science, 21*(12), 1854–1862.

Lyyra, P., Mäkelä, H., Hietanen, J. K., & Astikainen, P. (2014). Implicit binding of facial features during change blindness. *PLoS One, 9*(1), Article e87682.

MacDorman, K. F., Green, R. D., Ho, C. C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior, 25*(3), 695–710.

Mathur, V. A., Richeson, J. A., Paice, J. A., Muzyka, M., & Chiao, J. Y. (2014). Racial bias in pain perception and response: Experimental examination of automatic and deliberate processes. *The Journal of Pain, 15*(5), 476–484.

Mays, V. M., Cochran, S. D., & Barnes, N. W. (2007). Race, race-based discrimination, and health outcomes among African Americans. *Annual Review of Psychology, 58*, 201–225.

Meghani, S. H., Byun, E., & Gallagher, R. M. (2012). Time to take stock: A meta-analysis and systematic review of analgesic treatment disparities for pain in the United States. *Pain Medicine, 13*(2), 150–174.

Mende-Siedlecki, P., Lin, J., Ferron, S., Gibbons, C., Drain, A., & Goharzad, A. (2021). Seeing no pain: Assessing the generalizability of racial bias in pain perception. *Emotion, 21*(5), 932–950.

Mende-Siedlecki, P., Qu-Lee, J., Backer, R., & Van Bavel, J. J. (2019). Perceptual contributions to racial bias in pain recognition. *Journal of Experimental Psychology: General, 148*(5), 863.

Mende-Siedlecki, P., Qu-Lee, J., Lin, J., Drain, A., & Goharzad, A. (2020). The Delaware pain database: A set of painful expressions and corresponding norming data. *Pain Reports, 5*(6), Article e853.

Niv, D., & Kreitler, S. (2001). Pain and quality of life. *Pain Practice, 1*(2), 150–161.

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*(4), 498.

Oruc, I., Shafai, F., Murthy, S., Lages, P., & Ton, T. (2019). The adult face-diet: A naturalistic observation study. *Vision Research, 157*, 222–229.

Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology, 34*, 181–209.

Pasero, C., & McCaffery, M. (2010). *Pain assessment and pharmacologic management-E-book*. Elsevier Health Sciences.

Pauker, K., Rule, N. O., & Ambady, N. (2010). Ambiguity and social perception. In *The social psychology of visual perception* (pp. 13–29).

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13.

Penner, L. A., Hagiwara, N., Eggly, S., Gaertner, S. L., Albrecht, T. L., & Dovidio, J. F. (2013). Racial healthcare disparities: A social psychological analysis. *European Review of Social Psychology, 24*(1), 70–122.

Pesciarelli, F., Sarlo, M., & Leo, I. (2011). The time course of implicit processing of facial features: an event-related potential study. *Neuropsychologia, 49*(5), 1154–1161.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49*(1), 65–85.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*(3), 811.

Poole, G. D., & Craig, K. D. (1992). Judgments of genuine, suppressed, and faked facial expressions of pain. *Journal of Personality and Social Psychology, 63*(5), 797.

Prkachin, K. M., Berzins, S., & Mercer, S. R. (1994). Encoding and decoding of pain expressions: a judgement study. *Pain, 58*(2), 253–259.

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles, 59*(5–6), 377–391.

Reicherts, P., Wieser, M. J., Gerdes, A. B., Likowski, K. U., Weyers, P., Mühlberger, A., & Pauli, P. (2012). Electrocortical evidence for preferential processing of dynamic pain expressions compared to other emotional expressions. *PAIN®, 153*(9), 1959–1964.

Reskin, B. (2012). The race discrimination system. *Annual Review of Sociology, 38*, 17–35.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*(3), 499–505.

Richeson, J. A., Trawalter, S., & Shelton, J. N. (2005). African Americans' implicit racial attitudes and the depletion of executive function after interracial interactions. *Social Cognition, 23*(4), 336–352.

Ringwalt, C., Roberts, A. W., Gugelmann, H., & Skinner, A. C. (2015). Racial disparities across provider specialties in opioid prescriptions dispensed to Medicaid beneficiaries with chronic noncancer pain. *Pain Medicine, 16*(4), 633–640.

Schindler, S., Tirloni, C., Bruchmann, M., & Straube, T. (2021). Face and emotional expression processing under continuous perceptual load tasks: an ERP study. *Biological Psychology, 161*, 108056.

Shafto, J. P., & Pitts, M. A. (2015). Neural signatures of conscious face perception in an inattentional blindness paradigm. *Journal of Neuroscience, 35*(31), 10940–10948.

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *PAIN®, 135*(1–2), 55–64.

Suzuki, M., & Noguchi, Y. (2013). Reversal of the face-inversion effect in N170 under unconscious visual processing. *Neuropsychologia, 51*(3), 400–409.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813–833.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39*(6), 549–562.

Trawalter, S., Bart-Plange, D. J., & Hoffman, K. M. (2020). A socioecological psychology of racism: Making structures and history more visible. *Current Opinion in Psychology, 32*, 47–51.

Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial bias in perceptions of others' pain. *PLoS One, 7*(11), Article e48546.

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

Vervoort, T., Trost, Z., Prkachin, K. M., & Mueller, S. C. (2013). Attentional processing of other's facial display of pain: An eye tracking study. *PAIN®, 154*(6), 836–844.

Waytz, A., Hoffman, K. M., & Trawalter, S. (2015). A superhumanization bias in Whites' perceptions of Blacks. *Social Psychological and Personality Science, 6*(3), 352–359.

Wells, N., Pasero, C., & McCaffery, M. (2008). *Improving the quality of care through pain assessment and management. Patient safety and quality: An evidence-based handbook for nurses*.

Westbrook, J. I., Raban, M. Z., Walter, S. R., & Douglas, H. (2018). Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: A prospective, direct observation study. *BMJ Quality and Safety, 27*(8), 655–663.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598.

Winter, N., Burleigh, T., Kennedy, R., & Clifford, S. (2019). *A simplified protocol to screen out VPS and international respondents using Qualtrics*. Available at SSRN 3327274.

Xiao, Y., Coppin, G., & Van Bavel, J. (2016). Perceiving the world through group-colored glasses: A perceptual model of intergroup relations. *Psychological Inquiry, 27*(4), 255–274.

Yamada, M., & Decety, J. (2009). Unconscious affective processing and empathy: An investigation of subliminal priming on the detection of painful facial expressions. *Pain, 143*(1–2), 71–75.