Targets' facial width-to-height ratio biases pain judgments☆

Jason C. Deska*, Kurt Hugenberg

Miami University, 501 East High Street, Oxford, OH 45056, United States

ARTICLE INFO

ABSTRACT

The accurate perception of others' pain is important for both perceivers and targets. Yet, like other person perception judgments, pain judgments are prone to biases. Although past work has begun detailing characteristics of targets that can bias pain judgments (e.g., race, gender), the current work examines a novel source of bias inherent to all targets: structural characteristics of the human face. Specifically, we present four studies demonstrating that *facial width-to-height ratio*, a stable feature of all faces, biases pain judgments. Compared to those with low facial width-to-height ratio, individuals with high facial width-to-height ratio are perceived as experiencing less pain in otherwise identical situations (Studies 1, 2, & 3), and as needing less pain medication to salve their injuries (Study 4). This process was observed for White but not Black targets (Study 2), and manipulations of facial width-to-height ratio affected pain perceptions even when target identity was held constant (Study 4). Together, these findings implicate face structure in judgments of others' pain.

1. Introduction

Accurately perceiving others' pain is functionally important both for those experiencing suffering and for those judging others' pain. Understanding when others are in pain may be a key component of empathy (Cheng, Yang, Lin, Lee, & Decety, 2008) and can help perceivers distinguish between those who are in actual need of assistance from those who are not. Although extensive past work has explored processes by which perceivers can accurately detect others' pain (e.g., empathy), the current work focuses instead on questions about how *biases* rooted in person perception influence judgments of others' pain. Indeed, a variety of target-level characteristics, ranging from social categories such as race and gender, to embodied facial cues such as babyfacedness, can bias our judgments of others and their capacities (for a review, see Kawakami, Amodio, & Hugenberg, 2017).

In the current work, we extend this well-established research tradition investigating biases in person perception to investigate how targets' stable facial characteristics can bias how perceivers judge targets' experience of pain. Specifically, we investigate whether variations in a face structure inherent in all human faces – *facial width-to-height ratio* (fWHR) – biases judgments when determining the extent to which others experience pain. To that end, we first provide a brief summary of factors that bias pain judgments, with a particular focus on the role of target characteristics in pain perception biases. We then turn to a discussion about the role of fWHR in person perception and how fWHR may bias the perception of others' pain. Finally, we present four novel studies providing evidence for the hypothesized relationship between fWHR and pain perception.

2. Accuracy and bias in pain perception

Past work in the pain perception literature has primarily focused on the processes underlying pain perception *accuracy*. However, in the current work we pivot away from questions of accurate pain perception, instead focusing on characteristics inherent to the target that might *bias* pain judgments. Indeed, we believe that understanding how judgments of others' pain might be biased by perceivers' beliefs is an important topic of study. For example, patients expect that physicians can provide accurate diagnoses of pain, and society expects that jurors can render credible judgments of victims' harm based on their experiences of pain. However, when biases occur in such judgments, the interpretation of whether patients need treatment or whether victims need justice can be distorted. Research on biases in pain perception indicates that such biases can result from characteristics of the target, which we briefly review below.

One target characteristic from which pain judgment biases stem is gender. Research suggests that both men and women believe that men are less willing to report pain than women, and believe that women are more sensitive to pain and are less able to endure pain than men (Robinson et al., 2001). Meta-analytic evidence supports the notion that masculinity (which itself is associated with stoicism; Sanford, Kersh, Thorn, Rich, & Ward, 2002) is positively associated with both greater

pain tolerance and greater pain threshold (Alabas, Tashani, Tabasam, & Johnson, 2012). Experimental evidence corroborates this gender-based bias in pain perception. For instance, Robinson and Wise (2003) demonstrated that pain is underestimated for both men and women, but that this effect is larger for men. Relatedly, Pronina and Rule (2014) also found that pain is underestimated for both men and women; however, they found that participants primed with masculine concepts judged targets as experiencing less pain than when primed with feminine concepts.

Target race is another cue that generates robust and pervasive biases in pain judgments. Extensive evidence documents racial disparities in healthcare (Anderson, Green, & Payne, 2009; Green et al., 2003; Smedley, Stith, & Nelson, 2013). In particular, one domain in which healthcare reliably differs across race is in pain regulation. Compared to White individuals, people of color reliably have their pain under-estimated and undertreated (e.g., Anderson et al., 2009). One possibi-lity is that the pain experienced by people of color is recognized but under-treated, perhaps due to unequal access to healthcare or racial prejudice. However, recent research adopting a person perception fra-mework suggests that the pain experienced by people of color may simply not be recognized in the first place (Dore, Hoffman, Lillard, & Trawalter, 2014; Hoffman, Trawalter, Axt, & Oliver, 2016; Hollingshead, Meints, Miller, Robinson, & Hirsh, 2016; Trawalter, Hoffman, & Waytz, 2012; Waytz et al., 2015). Indeed, laypeople, chil-dren, nurses, medical students, and medical residents have all demon-strated beliefs that Black people feel less pain than White people in otherwise identical situations (Dore et al., 2014; Hoffman et al., 2016; Trawalter et al., 2012). Notably, these beliefs are independent of racial prejudice and negative racial attitudes (i.e., race-based antipathy). In-stead, the tendency to believe that Black people experience less pain than White people is associated with false beliefs linking Black people to non-human capabilities (Waytz et al., 2015) as well as erroneous beliefs about the biological differences between Black and White people (Hoffman et al., 2016).

Whereas the limited literature on biases in pain perception has fo-cused primarily on social categories (e.g., gender, race), in the present work, we seek to extend this understanding of biases in pain perception to facial width-to-height ratio. We turn now to a discussion of fWHR with a particular focus on how and why we hypothesized it may be implicated in pain perception.

3. Facial width-to-height ratio (fWHR)

Facial width-to-height ratio is a static face structure inherent to all faces. As shown in Fig. 1, fWHR is typically measured as the ratio be-tween bizygomatic width and upper face height (Hehman, Flake, & Freeman, 2015; Weston, Friday, & Liò, 2007). Although ori-ginally thought to be sexually dimorphic (Weston et al., 2007), mounting research finds little evidence for a strong relationship

between fWHR and sex (Geniole, Denson, Dixson, Carré, & McCormick, 2015; Kramer, 2017; Kramer, Jones, & Ward, 2012; Lefevre et al., 2012; Özener, 2012).

Research does, however, find abundant evidence linking fWHR with interpersonal aggression and behavioral dominance. For instance, high fWHR hockey players are more likely to spend time in the penalty box (Carré & McCormick, 2008). Indeed, meta-analytic evidence provides reliable support for the notion that relatively high fWHR individuals tend to be aggressive (Geniole et al., 2015; Haselhuhn, Ormiston, & Wong, 2015). In addition, fWHR is associated with how aggressive a person appears to be (e.g., high fWHR faces are rated as more likely to be guilty of assault than are low fWHR faces, Deska, Lloyd, & Hugenberg, 2017b). Recent work has demonstrated that both high fWHR men (Carré & McCormick, 2008; Carré, McCormick, & Mondloch, 2009) and women (Geniole, Keyes, Mondloch, Carré, & McCormick, 2012) are perceived as more ag-gressive than their low fWHR counterparts. Thus, fWHR both accurately predicts aggressive behavior and appears to be associated with a facial stereotype of behavioral dominance and aggression.

Building on these links between target fWHR and behavioral dom-inance and aggression, as well as past demonstrations that target level characteristics such as gender and race bias judgments of others' pain, we proposed that targets' fWHR might similarly bias pain judgments. Specifically, we hypothesized that high fWHR targets may be seen as experiencing relatively less pain than low fWHR targets, which may have downstream consequences such as perceiving less need for pain medication to salve their injuries. Insofar as high fWHR targets are seen as relatively brutish and animalistic (e.g., Deska et al., 2017b), they might also be seen as relatively insensitive to pain.

In the current work, we provide evidence across four studies for the hypothesis that perceivers would rate high fWHR individuals as ex-periencing less pain in otherwise identical situations and need less pain medication to salve their wounds, relative to low fWHR individuals. In Study 1, we sought to provide initial evidence for this hypothesis by showing participants individuals who naturalistically varied in their fWHR and asking them to assess how much pain they believed each target would feel across a variety of situations. Study 2 was designed as a replication and extension of Study 1. The design was identical except for the inclusion of both Black and White faces to test if race and fWHR independently or interactively bias pain judgments. A third study was designed to provide an additional replication and extension of the previous studies while also including a series of exploratory media-tional analyses to investigate potential mechanisms underlying a fWHR-based bias in pain judgments. Finally, Study 4 was designed to provide a more stringent test of the hypothesis that fWHR biases pain judg-ments. Rather than selecting faces that varied naturalistically on fWHR, we instead employed faces that were digitally manipulated to be higher or lower in fWHR, allowing us to hold target identity (and other non-fWHR parts of the face) constant while manipulating target fWHR.

Fig. 1. Facial width-to-height ratio is measured as the ratio between bizygomatic width (i.e., distance between left to right zygion) and upper face height (i.e., distance between mid-brow and upper lip). The face on the left has relatively low fWHR whereas the face on the right has relatively high fWHR.
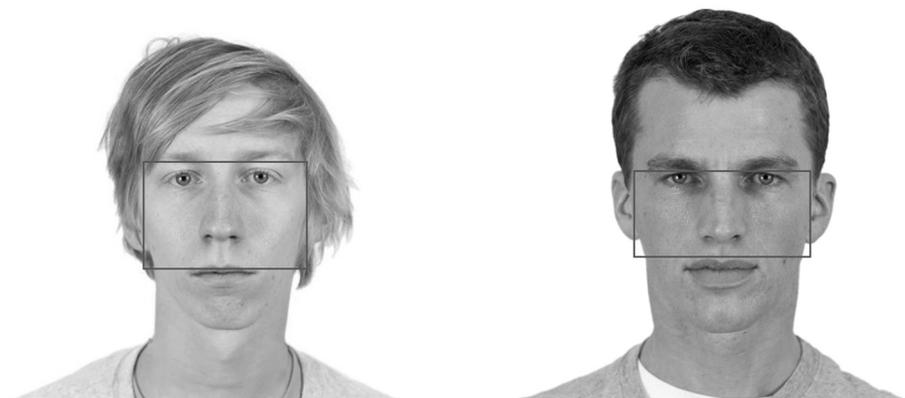
3.1. Study 1

Study 1 was designed to provide an initial test of the hypothesis that perceivers would believe that high fWHR individuals experience less pain than low fWHR individuals in otherwise identical situations. To test this hypothesis, participants were shown a series of faces that naturally varied in fWHR and were asked to assess how much pain the depicted individual would feel in a variety of potentially painful situations (Trawalter et al., 2012). We hypothesized that participants would believe that the high fWHR individuals would experience less pain than their low fWHR counterparts.

3.1.1. Method

3.1.1.1. Statistical power and participants.
Because we were uncertain as to the effect sizes we may observe in the current work, we relied on the sample size of the most analogous study in the literature, which was Trawalter et al.' (2012, Study 1) demonstration of race's effect on pain perception. We used their obtained effect size ($d = 0.51$) to estimate our sample size using G*Power (V3.1; Faul, Erdfelder, Lang, & Buchner, 2007). This analysis suggested we collect at least 63 participants to obtain 80% power. However, given that our independent variable was not race, but fWHR, we conservatively targeted at least 70 participants, and ran for as many complete weeks needed to achieve this sample size. Consequently, 96 undergraduate students ($M_{age} = 18.51$, $SD = 1.74$) participated in exchange for partial course credit. Most participants identified as White (80.4%) and female (83.5%).[1] No participants were excluded from analysis. All data were collected before analyses were conducted. All measures, manipulations, and exclusions are reported in these studies.

3.1.1.2. Materials.
Stimuli comprised of 20 images from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015). Specifically, we chose faces of the 10 White males who had the highest fWHR ($M = 2.04$, $SD = 0.06$) and faces of the 10 White men who had the lowest fWHR ($M = 1.74$, $SD = 0.03$). All faces displayed neutral expressions. It is worth noting that we held target race constant for our primarily White participant population because of the robust effects of race on both face perception (Hugenberg & Wilson, 2013) and pain perception (Trawalter et al., 2012). We held target gender constant because of past evidence that target gender biases pain perception (Robinson & Wise, 2003) and because most research with fWHR focuses on male targets. We directly test the effects of a race manipulation in Study 2. Pretesting (using normed data included with the Chicago Face Database) revealed that the high and low fWHR groups differed significantly on fWHR, $t(18) = 14.46$, $p < .001$, $d = 6.32$. However, high and low fWHR faces did not differ on attractiveness ($Ms = 2.18$ vs. 2.21, $SDs = 0.49$ vs. 0.62), $t(18) = -0.79$, $p = .44$, $d = -0.05$. Images were resized to 682×480 pixels and were presented in color.

3.1.1.3. Procedure.
Participants completed the study in individual

---

cubicles. All stimuli were presented via computer. After providing informed consent, participants were informed that past research has demonstrated that people are surprisingly accurate at making first impressions from faces and that the current research would investigate whether this extended to perceptions of pain tolerance. Participants then saw, one at a time, 20 faces (10 high fWHR, 10 low fWHR). Each face was presented simultaneously with 18 potentially painful scenarios taken directly from Trawalter et al. (2012). Example scenarios included disinfecting a sore and having one's fingers get caught in a door. For each scenario, participants were asked to estimate how much pain the depicted individual would experience. Pain ratings were made on 4-point scales ranging from 1 (*not painful*) to 4 (*extremely painful*). Both face presentation order and item presentation order (within face) were randomized.

After completing the pain ratings for a given target, participants were then presented with a new face and completed ratings of the same 18 scenarios for the newly depicted target. Presentation order of targets was independently randomized for each participant. Participants continued until ratings were made for each face. In total, participants made 360 ratings (i.e., 18 ratings for 20 faces). After completing the pain rating task, participants were asked to provide demographic information (e.g., gender, age, race), were thoroughly debriefed, and were thanked for their participation.

3.1.2. Results

Repeated-measures designs in which participants evaluate samples of stimuli can inflate Type 1 error rates when analyzed with ANOVAs (Judd, Westfall, & Kenny, 2012; Judd, Westfall, & Kenny, 2017). Consequently, in this and all studies, we used mixed-model analyses that fully crossed participants and stimuli and treated both as random factors. Facial width-to-height ratio was effect coded in all studies (low fWHR = -0.5, high fWHR = 0.5).

Of primary interest was the extent to which participants believed high and low fWHR targets would experience different amounts of pain across the same situations. Preliminary analyses indicated that the scenarios showed good internal reliability ($\alpha = .94$); we therefore computed mean pain ratings separately for the high and low fWHR faces. These pain ratings were then subjected to a cross-classified mixed-model analysis. Consistent with predictions, participants rated the high fWHR targets as experiencing less pain ($M = 1.85$, $SD = 0.39$) than the low fWHR targets ($M = 2.10$, $SD = 0.40$), $B = -0.25$, $SE = 0.07$, 95% CI $[-0.41, -0.09]$, $t(12.31) = -3.44$, $p = .005$, $d = -1.96$.

3.1.3. Discussion

Study 1 provided initial evidence that fWHR biases perceptions of others' pain. Specifically, participants believed that individuals with high fWHR would experience less pain than their low fWHR counterparts in otherwise identical situations. In Study 2, we sought to replicate and extend this effect by investigating whether similar effects would occur across target race. Specifically, in Study 2, we asked participants to make ratings of both White and Black male faces to test for the possible interactive role of fWHR and race in judgments of others' pain.

3.2. Study 2

Results from Study 1 demonstrated that fWHR biases perceptions of others' pain among a set of White male faces. We held target race constant for several reasons. First, race has robust effects on person perception and face perception (Hugenberg & Wilson, 2013). Second, past work has documented numerous race effects in interpersonal sensitivity judgments (e.g., (Elfenbein & Ambady, 2002; Hugenberg & Bodenhausen, 2003; Lloyd, Hugenberg, McConnell, Kunstman, & Deska, 2017). Thus, we thought it prudent to first demonstrate the existence of a fWHR-based bias using a racial

---

[1] In this and in all subsequent studies, we conducted preliminary, exploratory analyses to investigate whether participant gender had any effects and whether such effects qualified the predicted results. In Studies 1–3, no main effects or interactions including participant gender emerged (Study 1 main effect: $B = -0.04$, $SE = 0.08$, 95% CI $[-0.22, 0.13]$, $t(24.39) = -0.52$, $p = .607$, $d = -0.21$; Study 1 interaction: $B = 0.01$, $SE = 0.07$, 95% CI $[-0.14, 0.15]$, $t(13.00) = 0.09$, $p = .930$, $d = 0.05$; Study 2 main effect: $B = -0.09$, $SE = 0.08$, 95% CI $[-0.26, 0.08]$, $t(28.45) = -1.09$, $p = .286$, $d = -0.41$; Study 2 interaction: $B = 0.04$, $SE = 0.05$, 95% CI $[-0.07, 0.14]$, $t(17.92) = 0.69$, $p = .499$, $d = 0.33$; Study 3 main effect: $B = -0.08$, $SE = 0.07$, 95% CI $[-0.23, 0.06]$, $t(80.81) = -1.19$, $p = .239$, $d = -0.27$; Study 3 interaction: $B = -0.09$, $SE = 0.06$, 95% CI $[-0.20, 0.03]$, $t(51.36) = -1.44$, $p = .155$, $d = -0.40$). In Study 4, there was no main effect of gender ($B = -0.00$, $SE = 0.01$, 95% CI $[-0.20, 0.03]$, $t(84,916,396.81) = -0.00$, $p = .999$, $d = -0.00$), but there was a marginal interaction between participant gender and target fWHR, $B = 0.08$, $SE = 0.05$, 95% CI $[-0.01, 0.17]$, $t(228.98) = -1.73$, $p = .085$, $d = -0.23$. However, given that this was not observed in any other study, nor was it significant, it is likely attributable to chance and is not discussed further.
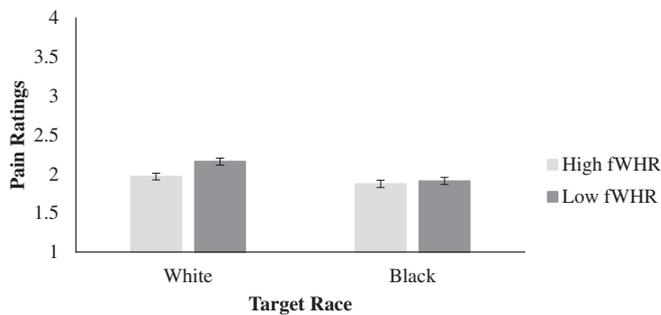
Fig. 2. Estimations of experienced pain as a function of target race and fWHR in Study 2. High fWHR targets were perceived as experiencing less pain than low fWHR targets, but only among White targets. Error bars represent standard error of the mean.

homogenous set of stimuli. However, as noted above, past work has extensively documented that race consistently biases judgments of others' pain (e.g., Hoffman et al., 2016; Trawalter et al., 2012; Waytz et al., 2015).

An open question was whether targets' fWHR would interact with target race in biasing perceptions of pain. Consequently, Study 2 had several goals. First, we sought to replicate the effect of fWHR on pain judgments among the White male faces used in Study 1. Additionally, the inclusion of Black male faces allowed us to provide further evidence for the existence of race-based biases in pain judgments and to test for interactive effects of race and fWHR. Here, two patterns of data seemed plausible. First, one sensible possibility was two main effects of fWHR and race, replicating Study 1 above as well as replicating the work of Trawalter, Hoffman, Waytz and colleagues who have consistently found a robust race effect on pain perception. Second, it is well established that racial outgroup faces are processed less configurally than are racial ingroups (Michel, Rossion, Han, Chung, & Caldara, 2006; Tanaka, Kiefer, & Bukach, 2004; Young & Hugenberg, 2010). Because fWHR is a configuration (i.e., a face ratio), the reduction in configural processing for racial outgroup faces could result in fWHR having a stronger effect on judgments of ingroup (i.e., White) than outgroup (i.e., Black) faces for our primarily White participant population. As a result, a second possibility was that target fWHR and target race might interact such that the effects of fWHR on pain perception are larger for White than for Black faces.

3.2.1. Method

3.2.1.1. Participants. We relied on the same sample size estimate and participant population for Study 2 as we did for Study 1. As before, we collected data for as many complete weeks needed to reach our target sample size and analyzed results only after data collection concluded. In Study 2, 75 undergraduate students ($M_{age} = 18.64$, $SD = 0.95$) participated in exchange for partial course credit. Most participants identified as White (80.0%) and female (66.7%). No participants were excluded from analysis.

3.2.1.2. Materials. We employed the same 20 White male faces in Study 2 that we used in Study 1 and we added 20 Black male faces to the current design as well. We used the same selection criteria for Black male faces as we did for the White male faces, sampling from the Chicago Face Database (Ma et al., 2015). Specifically, we chose 10 faces of Black males who had the highest fWHR ($M = 2.02$, $SD = 0.04$) and 10 faces of Black males who had the lowest fWHR ($M = 1.75$, $SD = 0.05$). Using the normed values from the Chicago Face Database revealed that these high and low fWHR groups differed significantly on fWHR, $t(18) = 12.89$, $p < .001$, $d = 2.88$, but that our high and low fWHR targets were not rated as differentially attractive across racial lines, $F(1,39) = 1.96$, $p = .17$, $\eta_p^2 = .05$. All images were resized to 682 × 480 pixels and were presented in color.

3.2.1.3. Procedure. The procedure for Study 2 was identical to that of Study 1 except as follows. Participants completed the 18-item pain scale for 20 total faces. Half of these faces were White and half were Black. Of the faces participants saw, half were high fWHR and half were low fWHR, which were distributed evenly across White and Black faces. Thus, participants saw five high-fWHR White, five high-fWHR Black, five low-fWHR White and five low-fWHR Black faces. Which set of faces participants viewed was counterbalanced. After completing the pain rating task, participants were asked to provide demographic information (e.g., gender, age, race), were thoroughly debriefed, and were thanked for their participation.

3.2.2. Results

Of primary interest was the extent to which target fWHR and race influenced participants' beliefs about the amount of pain an individual might experience in a series of hypothetical situations. Target race was effect coded (White = − 0.5, Black = 0.5). Replicating Study 1, this test yielded a main effect of target fWHR, $B = -0.11$, $SE = 0.03$, 95% CI [− 0.18, − 0.05], $t(32.18) = -3.38$, $p = .002$, $d = -1.19$. High fWHR targets ($M = 1.97$, $SD = 0.38$) were seen as experiencing less pain than were low fWHR targets ($M = 2.16$, $SD = 0.41$). Moreover, and consistent with past research, we also observed a main effect of target race, $B = -0.17$, $SE = 0.04$, 95% CI [− 0.25, − 0.10], $t(24.50) = -4.75$, $p < .001$, $d = -1.92$. Black targets ($M = 1.89$, $SD = 0.38$) were seen as experiencing less pain than were White targets ($M = 2.06$, $SD = 0.37$). However, as shown in Fig. 2, these lower order effects were qualified by a significant fWHR by target race interaction, $B = 0.15$, $SE = 0.06$, 95% CI [0.02, 0.28], $t(26.18) = 2.38$, $p = .025$, $d = 0.93$. To better understand the nature of this interaction, we looked at the effect of fWHR on judgments of pain separately for White and Black targets. Among the White faces, high fWHR targets ($M = 1.97$, $SD = 0.37$) were seen as experiencing less pain than were low fWHR targets ($M = 2.16$, $SD = 0.41$), $B = -0.19$, $SE = 0.05$, 95% CI [− 0.31, − 0.07], $t(8.14) = -3.56$, $p = .007$, $d = -2.50$. However, among the Black faces, high fWHR targets ($M = 1.87$, $SD = 0.39$) were seen as experiencing equal pain as the low fWHR targets ($M = 1.91$, $SD = 0.39$), $B = -0.04$, $SE = 0.04$, 95% CI [− 0.14, 0.06], $t(11.16) = -0.84$, $p = .418$, $d = -0.50$.

3.2.3. Discussion

The results of Study 2 replicated the findings of Study 1, demonstrating that high fWHR targets are believed to experience less pain than low fWHR targets. These results also replicated past work demonstrating that Black targets are believed to experience less pain that White targets (e.g., Hoffman et al., 2016; Trawalter et al., 2012; Waytz et al., 2015). However, we also observed a significant fWHR by target race interaction such that the effect of fWHR on pain judgments was significant for White targets but not for Black targets. The current study cannot specify why fWHR does not influence White perceivers' pain judgments of Black faces, however we think two possibilities are likely. One possibility is that our primary White-identified perceivers do not efficiently extract the width-to-height ratio from Black faces.[2] This is consistent with extensive past research demonstrating that racial outgroup faces are processed less configurally than are racial ingroup faces (e.g., Michel et al., 2006). Given that fWHR is inherently a face configuration (i.e., a relative difference between two different facial dimensions), there may be good reason to believe that the effects of fWHR may be particularly sensitive to factors that implicate configural processing, such as race. Another possibility is that perceivers are able to encode fWHR on Black faces, but do not use it. Rather, it may be the

---

[2] Restricting the sample to only White participants yields nearly identical effects (fWHR main effect: $B = -0.12$, $SE = 0.04$, 95% CI [− 0.19, − 0.05], $t(33.75) = -3.31$, $p = .002$, $d = -1.14$; race main effect: $B = -0.18$, $SE = 0.04$, 95% CI [− 0.26, − 0.10], $t(23.27) = -4.65$, $p < .001$, $d = -1.93$; interaction: $B = 0.15$, $SE = 0.07$, 95% CI [0.01, 0.29], $t(25.77) = 2.28$, $p = .031$, $d = 0.90$.

case that once perceivers categorize targets as Black, they rely solely on racial stereotypes (e.g., Hoffman et al., 2016; Trawalter et al., 2012; Waytz et al., 2015) when making pain judgments. Nevertheless, given that target race had a large effect on pain judgments and that fWHR did not influence ratings for Black targets, it is important to note the continued importance of studying the effects of target race on judgments of others' pain.

3.3. Study 3

The goals of Study 3 were twofold. First, Study 3 allowed us to provide an additional demonstration of the influence of fWHR on pain judgments. Additionally, in Study 3, we explored potential mechanisms for the fWHR-to-pain link. Although the first two studies provide evidence consistent with the hypothesis that fWHR biases perceivers' beliefs about the amount of pain others experience (at least for same-race White faces), they do not address the question of why this relationship exists. In particular, we tested the extent to which five theoretically motivated potential mechanisms (i.e., dehumanization, dominance, threat, anger, and masculinity) mediated the fWHR-to-pain link.

We measured dehumanization as a possible mediator because of recent work demonstrating that, compared to low fWHR targets, perceivers have a tendency to see high fWHR targets as less than fully human (Deska et al., 2017b). Thus, it may be the case that the tendency to dehumanize high fWHR targets (i.e., strip them of their sophisticated humanlike minds) might translate into a tendency to see them as less capable of experiencing pain. We measured dominance and threat because a large body of research finds that high fWHR targets both act and are perceived as dominant and threatening (e.g., Geniole et al., 2015). Similarly, research finds that high fWHR faces facilitate the recognition of, and share structural similarity to, anger (Deska, Lloyd, & Hugenberg, 2017a; Neth & Martinez, 2010). It seemed reasonable that formidable targets (i.e., targets that are angry, threatening, and dominant) might be judged as experiencing less pain compared to less formidable targets. Finally, we measured masculinity because past research has shown that masculinity is associated with perceived pain tolerance (e.g., Alabas et al., 2012). Although we had no specific a priori hypotheses about which potential mediators would yield significant indirect effects, each was theoretically motivated and thus plausible.

Consistent with Study 1, we showed participants faces of White men and asked them to judge how much pain they believed those individuals would experience across a variety of situations. Separately, participants were asked to complete a measure of blatant dehumanization for each face (Kteily, Bruneau, Waytz, & Cotterill, 2015). Finally, using the blatant dehumanization values as well as the pre-tested dominance, threat, anger, and masculinity values obtained from the Chicago Face Database (Ma et al., 2015), we use the Monte Carlo method (MacKinnon, Lockwood, & Williams, 2004) to test each of the 5 potential mediators in a mixed-model framework.

3.3.1. Method
3.3.1.1. *Participants*. We relied on the same sample size estimate for Study 3 as we did for the previous Studies, however in Study 3 we employed Mechanical Turk workers as participants. In Study 3, 84 mTurk workers ($M_{age} = 34.90$, $SD = 10.76$) participated and were remunerated $0.60. Most participants identified as White (73.8%) and male (52.4%). No participants were excluded from analysis.

3.3.1.2. *Materials*. Materials in Study 3 were identical to Study 1.

3.3.1.3. *Procedure*. The procedure for Study 3 was identical to that of Study 1 except as follows. In addition to completing the 18-item pain scale for all 20 faces, participants also completed a measure of blatant dehumanization for each face (Kteily et al., 2015). Participants were shown the classic "Ascent of Man" graphic with five silhouettes of a hominid varying in evolutionary stage. Participants were told that

"people can vary in how human-like they seem. Some people seem highly evolved whereas others seem no different from lower animals." Using a sliding scale paired with the "Ascent of Man" graphic, participants rated each target on how evolved they seemed. Participants made pain ratings and blatant dehumanization ratings in separate, counterbalanced, blocks. Counterbalancing had no effects on the results. Face order was independently randomized for each participant. After completing both tasks, participants were asked to provide demographic information (e.g., gender, age, race), were thoroughly debriefed, and were thanked for their participation. Compensation was not withheld from any participants.

3.3.2. Results
The first goal of the current study was to provide an additional demonstration of the effect of fWHR on pain judgments. Replicating the previous studies, high fWHR targets ($M = 2.24$, $SD = 0.37$) were perceived to experience less pain than low fWHR targets ($M = 2.43$, $SD = 0.33$), $B = -0.19$, $SE = 0.05$, 95% CI $[-0.29, -0.08]$, $t(27.50) = -3.55$, $p = .001$, $d = -1.35$.

Next, we tested the extent to which perceived dehumanization, dominance, threat, anger, and masculinity mediated the effect of fWHR on pain judgments. Values for dominance, threat, anger, and masculinity were obtained from the Chicago Face Database (Ma et al., 2015). To test mediation in a mixed-model framework, we used the Monte Carlo method (MacKinnon et al., 2004). For each analysis, we used 20,000 bootstrapped samples to test for a significant indirect effect. Although fWHR was significantly associated with each potential mediator (dehumanization: $B = 7.01$, $SE = 2.73$, 95% CI $[-12.66, -1.35]$, $t(22.97) = -2.56$, $p = .017$, $d = -1.07$; dominance: $B = 0.81$, $SE = 0.26$, 95% CI $[0.26, 1.36]$, $t(17.95) = 3.10$, $p = .006$, $d = 1.46$; threat: $B = 0.53$, $SE = 0.22$, 95% CI $[0.07, 0.99]$, $t(18.45) = 2.39$, $p = .027$, $d = 1.11$; anger: $B = 0.55$, $SE = 0.25$, 95% CI $[0.03, 1.07]$, $t(18.30) = 2.20$, $p = .041$, $d = 1.03$; masculinity: $B = 0.49$, $SE = 0.21$, 95% CI $[0.06, 0.93]$, $t(16.29) = 2.39$, $p = .029$, $d = 1.18$), results revealed that dehumanization, 95% CI $[-0.02, 0.01]$, perceived dominance, 95% CI $[-7.39, 11.26]$, perceived threat, 95% CI $[-0.19, 0.03]$, and perceived anger, 95% CI $[-0.18, 0.12]$, were not significant mediators. Results revealed that perceived masculinity, 95% CI $[-0.19, -0.01]$, was a significant mediator.

3.3.3. Discussion
The results of Study 3 provide an additional demonstration of the effect of fWHR on pain judgments, further demonstrating that individuals with high fWHR are assumed to experience less pain than their low fWHR counterparts. Additionally, these results help to provide some insight into the mechanism underlying the biasing effect of fWHR on pain judgments. Specifically, although we found that fWHR was significantly related to blatant dehumanization as well as perceived dominance, threat, anger, and masculinity, only perceived masculinity was a significant mediator. Thus, these results suggest that high fWHR faces appear more masculine than low fWHR faces, and this perceived masculinity in turn helps explain why perceivers judge high fWHR targets as experiencing less pain than their low fWHR counterparts. Although these results are interesting, we urge caution in making strong inferences from them. Although each potential mediator we tested was theoretically motivated by existing effects in the literature, the findings are exploratory in nature. Moreover, mediation tests cannot provide definitive evidence of causal relationships (e.g., Fiedler, Schott, & Meiser, 2011), nor can they distinguish between whether perceived masculinity (for example) mediates pain judgment effects, or whether pain judgment effects mediate masculinity judgments – statistically, both models are in the same equivalence class (Thoemmes, 2015). More work is needed to provide a more robust understanding of the mechanism(s) underlying the fWHR based bias in pain judgments. Nevertheless, the current results suggest that facial masculinity may have an important role, and are consistent with past work in the

literature linking pain judgments to targets' perceived masculinity (e.g., Alabas et al., 2012).

3.4. Study 4

To this point, each study relied on faces obtained from the Chicago Face Database (Ma et al., 2015) that varied naturalistically in whether they were high or low in fWHR. This design feature was intentional. The Chicago Face Database offers high quality images that are visually homogenous and tightly controlled. Nevertheless, a reasonable critique is that the previously observed effects might be unique to this set of faces. In Study 4, we manipulated fWHR within face identity. Holding target identity constant while manipulating the targets' fWHR provides a strong test of the hypothesis while allowing for within-target comparisons (i.e., a high fWHR as compared to a low fWHR version of the same face). Specifically, we employed previously created stimuli that hold face identity constant. Here, each face identity was manipulated to have a higher and lower fWHR version by digitally "stretching" or "squeezing" the stimuli (stimuli were taken from Stirrat & Perrett, 2012; see also Rowland & Perrett, 1995).

Additionally, in Study 4, we began to explore a downstream consequence of the previously observed biases: pain treatment. Whereas Studies 1–3 investigated how much pain perceivers believed targets would experience, in Study 4 we asked participants to judge which of two simultaneously presented targets would need more pain medication to treat otherwise identical injuries. Thus, in this design, we showed participants both digitally manipulated versions of each base face (i.e., the high and low fWHR version of the same identity) and asked participants to choose which of the two versions of the face would need more pain medication after experiencing a range of potentially painful outcomes. We hypothesized that participants would believe that high fWHR targets would need less pain medicine than low fWHR targets, even though identity was held constant.

3.4.1. Method
3.4.1.1. *Participants*. Because we were uncertain regarding the expected effect size due to the methodological change (i.e., switching to a forced-choice within-identity comparison), we conservatively opted to collect at least 200 participants. In actuality, 227 participants completed this study. Of these, 70 were undergraduate students (M_{age} = 18.67, SD = 0.93; 80.3% White; 64.3% female) and 157 were mTurk workers (M_{age} = 35.66, SD = 11.12; 67.5% White; 52.2% female). Because sample type did not significantly predict pain judgments, $t(225)$ = 1.44, p = .15, d = 0.19, nor did it interact with fWHR to predict pain judgments, B = −0.07, SE = 0.05, 95% CI [−0.17, 0.03], $t(195.93)$ = −1.45, p = .15, d = −0.21, we combined the two samples for the primary analysis. Undergraduate participants were compensated with partial course credit; mTurk workers were remunerated $0.50.

3.4.1.2. *Materials*. Stimuli consisted of 35 White male base faces (Stirrat & Perrett, 2012). Each base face was digitally manipulated into a low fWHR and a high fWHR version of the same base face as in past research (Rowland & Perrett, 1995; Stirrat & Perrett, 2012). This procedure yielded 70 total stimuli (i.e., 35 faces that were manipulated to be low fWHR and 35 faces that were manipulated to be high fWHR). All images were resized to 325 × 438 pixels and were presented in color.

3.4.1.3. *Procedure*. After providing informed consent, participants were informed that we were interested in perceptions of other people's pain tolerances. Participants completed 35 trials. During each trial, participants saw two faces: the high and low fWHR versions of a given face identity. Thus, each trial (presented in a randomized order) showed pairs of faces derived from one of the 35 total face identities. Below the pair of faces were seven items describing painful incidents

(e.g., involved in a car crash, had their wisdom teeth removed; Waytz et al., 2015). Participants rendered binary decisions about which target they believed would need more pain medicine after experiencing each of the seven painful incidents. These items showed strong internal reliability (α = .90). After completing the ratings, participants were asked to provide demographic information (e.g., gender, age, race), were thoroughly debriefed, and were thanked for their participation. Compensation was not withheld from any participants.

3.4.2. Results
Of primary interest was whether target fWHR influenced the amount of pain medication participants believed targets would need to ameliorate the pain they experienced after painful incidents. Counts of high and low fWHR selections were submitted to a mixed-model analysis along with participant gender (effect coded: male = −0.5, female = 0.5) and their interaction. This analysis yielded a marginally significant effect of fWHR on pain judgments, B = −0.05, SE = 0.03, 95% CI [−1.0, 0.01], $t(187.46)$ = −1.69, p = .094, d = −0.25. Converting the counts to proportions reveals that high fWHR faces (M = 0.48, SD = 0.17) were selected marginally less often than low fWHR faces (M = 0.52, SD = 0.17) when deciding which target would need more pain medication.

3.4.3. Discussion
Conceptually replicating the previous studies, the results of Study 4 provide further evidence for the biasing effect fWHR has on pain judgments. Specifically, in Study 4, participants tended to believed that high fWHR targets would need less pain medication following potentially painful incidents than low fWHR targets. This effect emerged on novel stimuli even when the actual identity of faces was held constant (i.e., fWHR was manipulated within-stimulus). However, given the effects were marginally significant, these inferences should be made with caution.

4. General discussion

Of interest in the current work was whether a stable facial characteristic – facial width-to-height ratio – would influence the perception of others' pain. We hypothesized that individuals with high fWHR would be perceived as experiencing less pain and need less pain medication to salve their wounds than individuals with low fWHR. In Study 1, we tested this hypothesis directly, showing participants faces that naturalistically varied on fWHR and asking them to report how much pain the depicted individuals might feel in a variety of potentially painful situations. Consistent with hypotheses, perceivers believed that high fWHR targets would experience less pain than low fWHR targets. In Study 2, we additionally manipulated target race, another feature of targets known to bias pain perception. Consistent with past work (e.g., Hoffman et al., 2016; Trawalter et al., 2012; Waytz et al., 2015), perceivers believed that Black targets would experience less pain than White targets. Additionally, replicating Study 1, high fWHR targets were seen as experiencing less pain than low fWHR targets. However, we also observed a target race by target fWHR interaction such that pain judgments were biased by fWHR for White targets but not Black targets. Study 3 explored potential mechanisms underlying the biasing effect of fWHR on pain judgment. Results of Study 3 provided another demonstration that high fWHR targets were seen as experiencing less pain than low fWHR targets and provided evidence that one potential mechanism through which this effect may occur is facial masculinity. Among our stimuli, higher fWHR men appeared more masculine which in turn predicted reductions in perceived pain (see also Pronina & Rule, 2014; Robinson & Wise, 2003). Finally, in Study 4, we manipulated fWHR within-target to provide a strong test of the fWHR hypothesis while holding constant other features of the faces. Consistent with the previous findings, participants tended to believe that the high fWHR targets would need less medication than the low fWHR targets, even
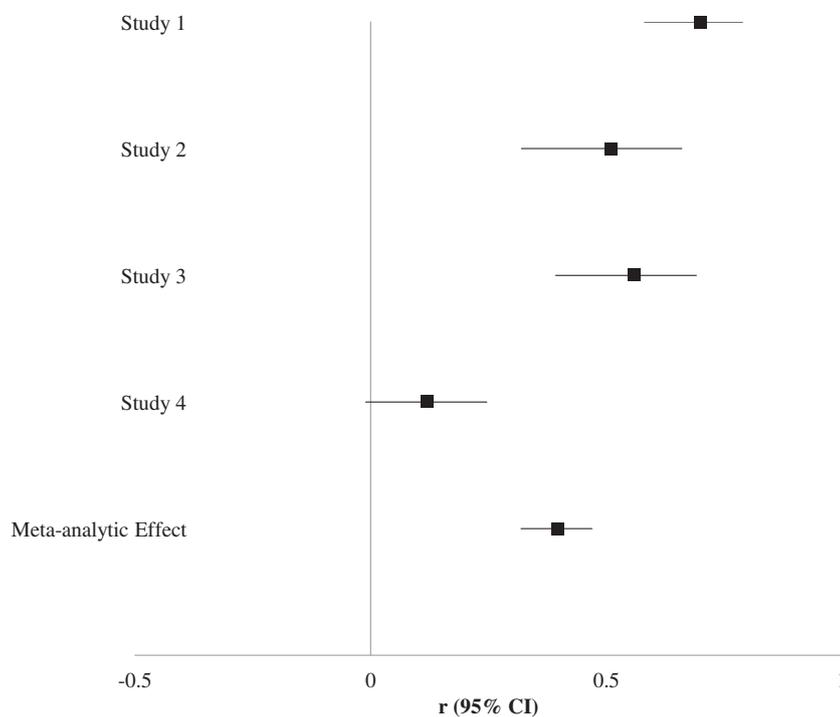
Study 1

Study 2

Study 3

Study 4

Meta-analytic Effect

-0.5 0 0.5 1

r (95% CI)

Fig. 3. Forest plot depicting the effect size *r* from each study as well as the overall meta-analytic effect. Error bars represent 95% confidence intervals around the effect size *r*.

with identity held constant.

To provide further evidence for the reliability of the effect of fWHR on pain judgments, we conducted a meta-analysis of our studies following procedures outlined by Goh, Hall, and Rosenthal (2016). To do this, we computed effect sizes for the main effects of fWHR in each study, yielding a total of four effects across 482 total participants. As shown in Fig. 3, this analysis yielded a statistically significant meta-analytic effect, $r = 0.40$, 95% CI [0.32, 0.47], $z = 9.11$, $p < .001$. The results of this meta-analysis provide cumulative evidence of the effect of fWHR on judgments of pain, suggesting that the effect is reliable and of medium magnitude by Cohen's conventions.

We believe that the current work has implications for a variety of literatures that are only infrequently linked. In particular, this research contributes specifically to work on facial width-to-height ratio as well as to the broader person perception literature. Previous research has documented a number of ways in which fWHR biases judgments of others. Individuals with relatively high fWHR are perceived to be behaviorally aggressive (Carré et al., 2009) intimidating (Hehman, Leitner, & Gaertner, 2013), dominant (Alrajih & Ward, 2014), threatening (Geniole et al., 2015), and less-than-fully human (Deska et al., 2017b). Here, we build on these existing findings to demonstrate a novel domain in which face-based stereotypes rooted in fWHR exert their influence.

Additionally, these findings have implications for research dedicated to improving the assessment and treatment of pain. Although work has begun documenting several areas in which characteristics of the perceiver (e.g., pain catastrophizing; Sullivan, Martel, Tripp, Savard, & Crombez, 2006) and target (e.g., race) influence judgments of others' pain, the current work adds to this literature by documenting a novel target characteristic that influences pain perception. All faces vary in facial width-to-height ratio; moreover, it is a static component of faces based in bone structure, not musculature. The current work demonstrates that this structural component of faces is sufficient to bias perceptions of how much pain an individual might be experiencing. Although preliminary, it is perhaps remarkable to consider that biases in pain judgment may occur for such a ubiquitous and often unnoticed face dimension.

4.1. Limitations and future directions

Despite the potential advances it offers, there are a number of limitations to the current work that we believe can provide exciting directions for future research. In particular, the participants in the current work were primarily undergraduates or were recruited online. It is unclear the extent to which these participants have similar biases in pain judgments as trained medical professionals. It may be the case that, with training, medical professionals are less susceptible to biases in their ratings of others' pain. Yet, extant work suggests this may not be the case. Work by Hoffman et al. (2016) reveals that medical professionals demonstrate the same race-based biases in pain perception as untrained perceivers. Nevertheless, the extent to which medical professionals are subject to the biases explored in the current work (i.e., face structure) remains an open question.

The current work also employed highly controlled and visually homogenous stimuli that were all displaying neutral expressions. As noted above, this design consideration was intentional. By using these stimuli, we can be reasonably confident that fWHR biases pain judgment, even when target identity is held constant (see Study 4). Nevertheless, it is unclear the extent to which similar biases might emerge with faces that are making veridical pain expressions. We hypothesize that fWHR would still exert an influence on pain judgments, such that even when all targets are displaying veridical pain expressions, participants would believe that the high fWHR targets are experiencing less pain than their low fWHR counterparts. People may even judge the expressions as less genuine on relatively high fWHR faces. Future research would do well to directly address this question.

Additionally, Studies 1–3 relied on faces obtained from the Chicago Face Database that varied quite drastically in their fWHR. This approach was intentional. The Chicago Face Database offers high quality, rigorously normed and controlled faces. Focusing on high and low fWHR targets allowed us to provide clear evidence for the existence of this bias in pain judgments. Nevertheless, this strategy may have caused our effects to be inflated. However, we showed analogous effects in Study 4 using a within-stimuli manipulation of fWHR. Moreover, the use of cross-classified analyses helps to limit potential concerns about stimulus effects. Ultimately, future work would do well to test the

extent to which this effect replicates with a variety of different types of stimuli and across a fuller range of fWHR.

In all studies, we employed only male faces as stimuli. This design choice was also intentional. Specifically, we employed only male stimuli to demonstrate the biasing effect of fWHR on pain judgments without the potentially confounding influence of target gender. Moreover, the majority of work examining the associations between fWHR and person perception does not investigate female faces (e.g., Geniole et al., 2015). Nevertheless, female faces also vary in fWHR and it is an open question as to whether the effects obtained with male faces generalize to female faces. It may be the case that target gender, which by itself biases pain perception (Robinson & Wise, 2003), interacts with target fWHR to influence pain judgments. Alternatively, it may be the case that fWHR has similar effects on female faces as it does on male faces (see Deska et al., 2017b). Future work is needed to better understand how fWHR interacts with a range of person variables to influence pain judgments, including target gender and target race, among others.

Additionally, the first three studies asked participants to report how much pain they believed targets would feel in hypothetical situations whereas the fourth study asked participants to report how much medication they believed targets would need to soothe their pain. Although we see both of these questions as important pain judgments, it could be argued that they are conceptually distinct. Whereas the former is tapping into pain intensity, the latter might be more closely assessing pain tolerance. We observed similar findings regardless of the specific form of pain judgment participants were asked to make, suggesting that the observed effects may generalize across domain. Nevertheless, future research would do well to directly assess biases in pain threshold, intensity, and tolerance.

Finally, it is unclear from the current work whether the biases in pain judgments reflect reality – perhaps high fWHR men actually can tolerate more pain than their low fWHR counterparts. In the current work, we opted to focus exclusively on whether target fWHR influences perceivers' pain judgments. An interesting possibility is that this facial stereotype linking fWHR to pain tolerance may have a kernel of truth. High fWHR targets may not only be seen as higher in pain tolerance, perhaps they actually *are* more tolerant of pain. Future work is needed to determine the extent to which fWHR correlates with actual pain tolerance and whether the judgment biases we observed here reflect an actual difference between higher and lower fWHR individuals.

4.2. Conclusion

In sum, the current work finds that facial width-to-height ratio biases pain judgments. People believe that high fWHR individuals, compared to low fWHR individuals, experience less pain in otherwise identical situations, and need less pain medication to salve their injuries, even when identity is held constant. The current work demonstrates that features of others' faces – and even often unnoticed features – can play a key role in how we simulate their experience of pain.

Open practices

All of the data presented in this manuscript are openly available online at https://osf.io/xzp9s. There is sufficient information for an independent researcher to reproduce the reported results.

References

Alabas, O. A., Tashani, O. A., Tabasam, G., & Johnson, M. I. (2012). Gender role affects experimental pain responses: A systematic review with meta-analysis. *European Journal of Pain, 16*, 1211–1223.

Alrajih, S., & Ward, J. (2014). Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *British Journal of Psychology, 105*, 153–161.

Anderson, K. O., Green, C. R., & Payne, R. (2009). Racial and ethnic disparities in pain: Causes and consequences of unequal care. *The Journal of Pain, 10*, 1187–1204.

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London B: Biological Sciences, 275*, 2651–2656.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20*, 1194–1198.

Cheng, Y., Yang, C. Y., Lin, C. P., Lee, P. L., & Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: A magnetoencephalography study. *NeuroImage, 40*, 1833–1840.

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2017a). The face of fear and anger: Facial width-to-height ratio biases recognition of angry and fearful expressions. *Emotion* (Advance online publication).

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2017b). Facing humanness: Facial width-to-height ratio predicts ascriptions of humanity. *Journal of Personality and Social Psychology* (Advance online publication).

Dore, R. A., Hoffman, K. M., Lillard, A. S., & Trawalter, S. (2014). Children's racial bias in perceptions of others' pain. *British Journal of Developmental Psychology, 32*, 218–231.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*, 203–235.

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

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analyses can (not) do. *Journal of Experimental Social Psychology, 47*, 1231–1236.

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PloS One, 10*, e0132726.

Geniole, S. N., Keyes, A. E., Mondloch, C. J., Carré, J. M., & McCormick, C. M. (2012). Facing aggression: Cues differ for female versus male faces. *PloS One, 7*, e30366.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*, 535–549.

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

Haselhuhn, M. P., Ormiston, M. E., & Wong, E. M. (2015). Men's facial width-to-height ratio predicts aggression: A meta-analysis. *PloS One, 10*, e0122637.

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin, 41*, 1123–1134.

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology, 49*, 747–754.

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*, 4296–4301.

Hollingshead, N. A., Meints, S. M., Miller, M. M., Robinson, M. E., & Hirsh, A. T. (2016). A comparison of race-related pain stereotypes held by white and black individuals. *Journal of Applied Social Psychology, 46*, 718–723.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640–643.

Hugenberg, K., & Wilson, J. P. (2013). Faces are central to social cognition. In D. Carlston (Ed.). *Handbook of social cognition* (pp. 167–193). New York, NY: Oxford University Press.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54–69.

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology, 68*, 601–625.

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology, 55*, 1–80.

Kramer, R. S. (2017). Sexual dimorphism of facial width-to-height ratio in human skulls and faces: A meta-analytical approach. *Evolution and Human Behavior, 38*, 414–420.

Kramer, R. S., Jones, A. L., & Ward, R. (2012). A lack of sexual dimorphism in width-to-height ratio in white European faces using 2D photographs, 3D scans, and anthropometry. *PloS One, 7*, e42705.

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*, 901–931.

Lefevre, C. E., Lewis, G. J., Bates, T. C., Dzhelyova, M., Coetzee, V., Deary, I. J., & Perrett, D. I. (2012). No evidence for sexual dimorphism of facial width-to-height ratio in four large adult samples. *Evolution and Human Behavior, 33*, 623–627.

Lloyd, E. P., Hugenberg, K., McConnell, A. R., Kunstman, J. W., & Deska, J. C. (2017). Black and white lies: Race based biases in deception detection. *Psychological Science, 28*, 1125–1136.

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

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99–128.

Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17*, 608–615.

Neth, D., & Martinez, A. M. (2010). A computational shape-based model of anger and sadness justifies a configural representation of faces. *Vision Research, 50*, 1693–1711.

Özener, B. (2012). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior, 33*, 169–173.

Pronina, I., & Rule, N. O. (2014). Inducing bias modulates sensitivity to nonverbal cues of others' pain. *European Journal of Pain, 18*, 1452–1457.

Robinson, M. E., Riley, J. L., Myers, C. D., Papas, R. K., Wise, E. A., Waxenberg, L. B., & Fillingim, R. B. (2001). Gender role expectations of pain: Relationship to sex differences in pain. *The Journal of Pain, 2*, 251–257.

Robinson, M. E., & Wise, E. A. (2003). Gender bias in the observation of experimental pain. *Pain, 104*, 259–264.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications, 15*, 70–76.

Sanford, S. D., Kersh, B. C., Thorn, B. E., Rich, M. A., & Ward, L. C. (2002). Psychosocial mediators of sex differences in pain responsivity. *The Journal of Pain, 3*, 58–64.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (2013). *Unequal treatment: Confronting racial and ethnic disparities in health care.* Washington, DC: National Academies Press.

Stirrat, M., & Perrett, D. I. (2012). Face structure predicts cooperation: Men with wider faces are more generous to their in-group when out-group competition is salient. *Psychological Science, 23*, 718–722.

Sullivan, M. J. L., Martel, M. O., Tripp, D. A., Savard, A., & Crombez, G. (2006). Catastrophic thinking and heightened perception of pain in others. *Pain, 123*, 37–44.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition, 93*, B1–B9.

Thoemmes, F. (2015). Reversing arrows in mediation models does not distinguish plausible models. *Basic and Applied Social Psychology, 37*, 226–234.

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

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

Weston, E. M., Friday, A. E., & Liò, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *PloS One, 2*, e710.

Young, S. G., & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology, 99*, 964–977.