On the Psychology of Gendered Colorism

Diane-Jo Klorkor Bart-Plange
Charlottesville, VA, USA

M.A. University of Virginia, 2019
B.A. Washington University in St. Louis, 2015

A Dissertation Presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy
Department of Psychology
University of Virginia
April 2022

Abstract

From structural inequalities to interpersonal bias, racism persists in American society. Colorism, a phenomenon borne out of racism, is a form of discrimination that creates widespread disparities between lighter- and darker-skinned Black people (Monk, 2015). Colorism is often gendered; in some instances, darker-skinned Black women and girls incur greater penalties for their skin tone than do darker-skinned Black men and boys (Hunter, 2002; Hunter, 2005; Alter et al., 2016; Hannon et al., 2013). The present work investigates when and why colorism becomes gendered.

In three studies, I examine the psychological underpinnings of colorism and gendered colorism. Specifically, I tested the following predictions: 1) Skin tone bias: Skin tone biases evaluations of Black men and women. 2) Gendered skin tone bias: Skin tone will bias evaluations of Black women more than Black men, particularly in domains where gender is salient. Studies 1 and 2 provided evidence of skin tone bias and gendered colorism in categorization and evaluation of prototypicality of race and gender. Study 3 extended these findings to a consequential area of healthcare for Black Americans—pain perception. Although I did not find direct evidence of skin tone bias in pain perception, skin tone influenced perceptions of race prototypicality, which in turn influenced assumptions of life hardship and ratings of pain. Moreover, for gendered pain scenarios, perceptions of gender prototypicality influenced assumptions of hardship and ratings of pain, particularly for Black women. Taken together, this work reveals how colorism manifests in individual cognition and provides avenues for future research.

Table of Contents

Dedication

This dissertation is dedicated to my elder sisters, Jayne Bart-Plange and Dr. Emma-Lorraine Bart-Plange, the embodiment of the phrase “phenomenal woman”. You are my biggest supporters and my best friends. Your love, kindness, and support through the past six years have been invaluable to me. This is for you.

Acknowledgements

I am grateful beyond words for so many people throughout this journey to earn my Ph.D. First, I would like to thank Sophie Trawalter, my primary advisor. I am incredibly fortunate to have had you as a mentor and to have your guidance. You have made me a better researcher. Most of all, you have shown me the kind of mentor I hope to be to others. You care not only about helping us grow academically, but about our well-being and growth as people with a duty to care for others. Thank you for all that you've taught me these past six years.

I'd also like to thank the rest of my dissertation committee—Jazmin Brown-Iannuzzi, Seanna Leath, and Gabrielle Adams. Thank you for your feedback, your mentorship, your kindness, and the inspiration you've given me to become a better researcher and critical thinker. Thank you so much for making this milestone possible.

Thank you to Jin Bak and Kyshia Henderson, my lab sisters. I am so grateful to know you, to learn from you, and to call you my friends. Thank you to all the friends and lovely people in my life who have shown me so much support from near and far and never let me forget that I am capable. Thank you for cheering me on every step of the way. Thank you to Nick Buttrick and Erin Westgate, who I jokingly, but in all seriousness, call my secondary advisors. I have learned so much from you and could not have imagined my graduate school experience without your guidance. You make research fun. Thank you to the many SCABL lab RAs who have helped make this research possible. Thank you to Nao Hagiwara and Danyel Smith, my collaborators at Virginia Commonwealth University who helped collect data for my final study.

Lastly, I'd like to thank my wonderful and loving family. You are my everything. I am blessed to have been raised by Joseph and Emma Bart-Plange. Every achievement I have was made possible because of you.

On The Psychology of Gendered Colorism

50+ years since the passage of the Civil Rights Act, racial disparities and discrimination stubbornly persist in American society. White Americans continue to have advantages while Black Americans lag behind in income, wealth, educational quality and achievement outcomes, physical health, treatment within healthcare, and more. In other words: racism is alive and well, and the impact of legally sanctioned discrimination has not been repaired. Born out of this existing system of white supremacy comes colorism, a form of discrimination that targets darker-skinned individuals and privileges light-skinned individuals in a racial group (Hunter, 2007).

Although comparatively less studied within psychology as racial prejudice and discrimination, colorism has significant impacts for darker-skinned Black Americans today in a myriad of areas of life; compared to lighter-skinned Black people, darker-skinned Black people report experiencing discrimination more often (Monk, 2015; Uzogara & Jackson, 2016), are arrested more often and receive harsher punitive sentences for similar offenses (Blair, Judd, & Chapleau, 2004; Eberhardt et al., 2006), have less occupational attainment, lower socioeconomic status, and more negative media representation (Hughes & Hertel, 1990; Blair, Judd, Sadler, & Jenkins, 2002; Maddox & Chase, 2004, Hunter, 2007; Alter et al., 2016). Some research suggests that disparities such as health outcomes between lighter- and darker-skinned Black people are just as large, if not larger than existing inequities between Black and White Americans (Monk, 2015), and further research has shown that people have and are aware of different stereotypes for lighter-skinned and darker-skinned Black people (Maddox & Gray, 2002). Colorism, then, is a necessary aspect of understanding the nature of racist discrimination, and an important factor to analyze in experimental work about it. In this dissertation, I examine the psychological underpinnings of colorism. I consider how historical context and current sociocultural dynamics

give rise to colorism and how colorism becomes gendered. I show with this work how colorism is an important albeit often neglected form of white supremacy and has particular consequences for Black women.

Historical Background and Colorism Today

In the context of the United States, colorism was borne from the system of slavery. Enslavers preferred enslaved Africans with more Eurocentric appearances, such as having lighter skin, straight and longer hair, and Eurocentric facial features (i.e., narrow nose, smaller lips; Reece, 2018; Russell et al., 1992), because these features indicated closer physical association to whiteness. Due to this, enslaved lighter-skinned Africans were often able to receive privileges such as higher likelihood for manumission and receiving any education, and lighter-skinned free Black Americans had greater opportunities for education and employment (Hughes & Hertel, 1990; Reece, 2018). Lighter-skinned Black people were also preferred as personal servants and household workers of their enslavers, while darker-skinned Black enslaved people were more often relegated to more grueling physical field labor outdoors (Reuter, 1918; Ryabov 2013). Lighter-skinned Black people undoubtedly were subjected to the horrors of slavery and racist discrimination; however, their relative higher societal standing in comparison to their darker-skinned counterparts led to long-lasting economic and social advantages over them. The impact of these manufactured hierarchies continues today: generations of systemic discrimination have led to present day intra-racial skin tone disparities in areas such as income and employment (Goldsmith et al., 2007), and lighter skin and Eurocentric features are still perceived as more attractive in American society (Hall, 2017; Hunter, 2007; Reece, 2018).

Colorism is often gendered. Patriarchal norms elevate attractiveness as social capital for women, having relatively more impact on the social status and economic outcomes of women

than men (Hunter, 2002; Hunter, 2007). Beauty and femininity are racialized concepts; the ubiquity of Eurocentric beauty norms prizing White people as the ultimate standard of beauty make being perceived as beautiful and feminine of greater importance for how Black women are perceived and treated by others.

Concerning colorism in the United States, historically, this too can be tied back to the racist and sexist treatment of Black women during slavery. Black feminist scholar belle hooks details in her 1981 book *Ain't I a Woman* that while Black men were generally forced to perform physical labor such as fieldwork, "the black female was exploited as a laborer in the fields, a worker in the domestic household, a breeder, and as an object of white male sexual assault" (p. 21). White society attributed dehumanizing and objectifying stereotypes to Black women to justify ***both*** their enslavement and sexual exploitation. These "controlling images" (Collins, 1990) of Black women were and continue to be closely tied to skin tone. The "Jezebel" is the image of a lascivious seductress and manipulative "mulatto" (mixed-race) Black woman with Eurocentric physical features, who cannot control her sexual desires, while the "Mammy" image is of a large, darker-skinned, asexual Black women with Afrocentric features, who lives to serve her "master" and serve as the nanny to his children (Townsend, 2010; Collins, 1990). These historical images reflect the gendered nature of colorism; the association of light skin to desirability and beauty has larger impacts on the perceptions of Black women, and the association of dark skin to both sexlessness and masculinity is specific to the historical perceptions of Black women.

Research in the past decade has begun to further investigate contexts where colorism creates harsher consequences for Black women and girls in the present era. An analysis of The National Longitudinal Study of Adolescent to Adult Health data showed that darker-skinned girls

are 3.4 times more likely to be suspended than lighter-skinned girls, but darker-skinned boys are only 2.5 times more likely to be suspended compared to their lighter-skinned counterparts (Hannon et al., 2013). In another analysis of the same dataset, Blake and colleagues (2017) found that darker-skinned Black girls are more likely than White girls to be suspended, but there is no significant difference between suspension rates of lighter-skinned Black girls and White girls. These harsher consequences for Black women can also be found in magazine depictions of celebrities. Alter and colleagues (2016) found in their analysis of pop culture magazines that the more negative the news story, the darker Black women were depicted, however, Black men's skin tone depictions did not differ based on news story content (Alter et al., 2016). In the present research, I take a bottom-up approach to understanding why colorism becomes gendered. I examine the cognitive mechanism behind colorism, how gendered contexts create this bias, and why it can so greatly impact the lives of Black women and girls.

On the Psychology of Colorism

The lengthy history of racism has had a great impact on shaping cultural narratives, norms, and perceptions of race and gender. Race and gender stereotypes play an important role in how we categorize others; for instance, research shows that sex categorization is biased by race, and vice versa. Black men are perceived as prototypical of the racial category "Black" due to shared stereotype content: Black people and men are both stereotyped as threatening, more aggressive, athletic, and dominant. "Blackness" is associated with masculinity, and this association aids Black men's racial categorization (Goff, Thomas, & Jackson, 2008). On the other hand, "woman" and "Black" categories have opposing stereotype content: women are stereotyped as friendly, kind, soft-spoken, and gentle (L. K. Johnson et al., 2012). Studies have found that people are slower to categorize Black women as "Black" compared to Black men, and

slower to categorize Black women as "women" compared to White women" (Cooley, et al., 2018; Thomas et al., 2014; L. K. Johnson, et al., 2012). The association of "Blackness" to maleness leads observers to perceive Black women as more masculine than White women, and therefore less prototypical of women (Goff, Thomas, & Jackson, 2008).

This non-prototypical status also has social consequences for Black women. The Intersectional Invisibility Hypothesis posits that intersectional targets are non-prototypical within their race and gender groups, and due to this, Black women often experience discrimination in forms of neglect or invisibility (Purdie-Vaughns & Eibach, 2008). Existing experimental research bears this out; studies find that Black women often go "unseen and unheard". Perceivers are less able to distinguish between Black women they have seen before compared to White men and women and Black men, and their contributions are more likely to be mis-remembered or attributed to others (Sesko & Biernat, 2010). Additional research provides evidence that gender saliency uniquely affects Black women, expanding this model of intersectionality invisibility. Cooley et al., (2018) showed that when presented with neutral faces of Black and White women, participants were quicker to recognize White women as women. However, when presented with smiling Black women, activating gender stereotypes of women being warm and friendly, participants were significantly quicker to categorize them as women compared to White women. This research shows that although Black women's perceived non-prototypicality leads to social invisibility, the salience of gender identity and associated stereotypes in any given context can lead to hyper-visibility, providing evidence that observers attend and respond to features that highlight their gender identity. This work on stereotyping and prototypicality gives us insight to how colorism becomes gendered on a cognitive level. Categorization and labeling of non-prototypical people is relatively more effortful compared to prototypical group members. I

propose that because of this, people may pay more attention to readily available features that provide categorization information for non-prototypical people. The perceived non-prototypicality of Black women for their race and gender leads perceivers to pay more attention to race stereotypic features, such as skin tone, when evaluating them. Due to the heavy association of Blackness with masculinity, this race stereotypic feature also signals gender—making skin tone a gendered feature as well. Further, I suggest that this bias happens particularly in gendered domains, which make the perceived non-prototypicality of Black women most salient.

In this dissertation, I connect these biases in categorization to evaluative judgments. I extend this theory to a consequential area of healthcare. Healthcare is a realm in which racial disparities are well-documented and color disparities research is increasing. Some of the existing literature thus far shows that darker-skinned Black people have worse self-reported physical and mental health, are much more likely to develop hypertension (Laidley et al, 2019; Ugozara & Jackson, 2016), and have higher rates of mortality among Black people with a high school or higher-level education (Stewart et al., 2018). Within the realm of healthcare, I focus on pain perception and management—an area that is a contributing factor to health disparities for Black Americans, important for the quality and efficiency of care in the medical system.

Years of research have uncovered striking racial inequities specifically in pain management. This bias is found even against young children: physicians have been found to be less likely to prescribe the appropriate treatment for severe pain (opioids) for Black children (vs. White) diagnosed with appendicitis (Goyal et al., 2015), and pediatricians with negative beliefs about Black patients' medical compliance are less likely to agree to prescribing recommended medication for Black patients with urinary tract infections (Sabin & Greenwald, 2012).

Researchers have also found this bias across various medical issues, such as metastatic cancer and extremity fractures (Cleeland et al., 1997; Todd et al., 2000). This has clear consequences for the quality of treatment and downstream effects such as trust in the medical system and likelihood of seeking medical care. Existing research shows that race and gender stereotypes and beliefs stemming from them, such as beliefs in biological differences between Black and White people and perceptions of life hardship, all have influence on how laypeople and medical professionals perceive and treat the pain of others (Samulowitz et al., 2018; Druckman et al., 2018; K. M. Hoffman et al., 2016). Given how much skin tone influences the stereotypes attributed to people within racial groups, as well as perceptions of their racial and gender prototypicality, skin tone likely has a substantial impact on pain perception for Black Americans, and unique consequences for Black women.

Where Gender, Skin tone, and Pain Bias Intersect

Current literature in pain perception management research shows that both gender and racial stereotypes play a role in how pain is perceived and treated in healthcare. Research shows that although women have more pain compared to men and it is more socially accepted for them to express it, women have their reports of pain taken less seriously, and are more likely to have their pain attributed to emotional factors regardless of the cause (D. E. Hoffmann & Tarzian, 2001; Samulowitz et al., 2018). Women reporting pain are often seen as hysterical, faking their pain, or needing to prove they are as ill as their male counterparts before being treated similarly (Barsky et al., 2001; D. E. Hoffmann & Tarzian, 2001; P. A. Johnson et al., 1996). A recent study on medical records of Black and White patients found that physicians have more linguistic bias for White and Black women, and Black men, compared to White men, indicating that they may not believe these patients' complaints more often (Beach et al., 2021). Thus, both racial (as

noted in the previous section) and gender disparities in patient treatment when they report health issues are prevalent in healthcare.

Within the many studies about how women's pain is treated in healthcare settings, most studies do not pay specific focus to how Black women fare when gendered norms and expectations have clear influences on treatment in healthcare. For instance, women's pain is often "psychologized", more likely to be attributed to emotional rather than somatic causes, and in general they are more likely to be prescribed antidepressants and are diagnosed more frequently with affective disorders (Samulowitz et al., 2018). However, when turning the focus to Black women, research shows that they may be both underdiagnosed and undertreated for affective disorders (Holden et al., 2013). More recent research focusing on Black women's treatment in healthcare presents a similar picture of undertreatment or maltreatment. A recent study showed that 41% of Black women report their women's health concerns not being taken seriously by medical providers (Funk, 2022). Overall, 50% of Black women over 50 and 70% of Black women between the ages of 18-49 have reported experiencing at least one type of negative experience (i.e. not having their pain taken seriously by a provider, being treated with less respect compared to other patients, having to speak up to receive proper care) while interacting with a health provider, and Black women overall are more likely to have had a range of negative healthcare experiences compared to Black men (Funk, 2022).

Another recent study of clinical vignettes found that while Black patients overall were 3.6% more likely to be refused pain treatment compared to White patients, a closer look revealed that Black women specifically were most likely to be undertreated, being 7% more likely to be denied pain treatment (Logé et al., 2021). As Black women face unique circumstances of invisibility or hypervisibility, or even additional maltreatment due to their intersecting identities,

examining how gender norms and expectations influence their treatment is a necessity. In the present work, I hypothesize that laypeople's perceptions of Black men and women's pain will be influenced by skin tone, and particularly so for Black women in contexts where gender is salient.

Overview of Studies

I use both the terms colorism and skin tone bias interchangeably for the purposes of this research but recognize that colorism goes beyond bias and discrimination at the individual and interpersonal levels alone, including systemic factors and cultural norms. In Studies 1a & b, I examined whether skin tone biases racial and gender categorization of Black men and women, and whether these biases are more pronounced for Black women. This study allowed me to test whether Black women are seen as less prototypical of their racial group and dark-skinned Black women as less prototypical of their gender group. These studies were conducted as part of my pre-dissertation work (see Bart-Plange, 2019), but includes new contrast analyses. I then build on Study 1 in two new studies. In Study 2, I investigate whether skin tone biases person perception—the mental processes that people use to form impressions of others—and specifically, the extent to which skin tone guides person perception. I also directly measure the extent to which dark- and light-skinned Black men and women are seen as prototypical of their gender and racial group, to document prototypicality biases. In Study 3, I extend this research and explored downstream consequences of skin tone bias in person perception. Here, I test whether skin tone influences laypeople's perceptions of Black men and women's pain.

Study 1a and b: Gender Salience Increases Skin Tone Bias in Categorization

Studies 1a and b were conducted as part of my pre-dissertation work. The aim of these studies was to expand past racial categorization research by testing how skin tone influences participants' racial and gender categorizations of Black men and women. In this study, I

measured participants' reaction time and accuracy of racial and gender categorizations of Black and White men and women. Because skin tone is the most visible and easily used characteristic associated with race (Feliciano, 2016) and the social constructions of race and gender are interrelated, I predicted that skin tone should consistently influence participants' ability to categorize by race and gender quickly and accurately. Beginning with basic categorization is a necessary starting point in examining how skin tone influences how prototypical Black people, especially women, are perceived socially. Stereotyping, prejudice, and discrimination stem from the cognitive processes of social categorization. Measuring reaction time and accuracy gives insight to how stereotypical Black men and women are imagined as members of their race or gender. If Black women are perceived as non-prototypical of women, when confidence is low, reaction time should increase, as more evidence (i.e., longer looking time) is needed before making an accurate categorization decision. In tasks where categories are well-defined (though the gender binary is a social construction itself, male and female genders are one of the first social categories people become aware of in life), inaccurate categorizations indicate not clearly knowing the rules of categorization. More errors and slow categorizations should occur for Black women if they are not perceived as a prototype for women, while Black men (particularly with darker skin) should be categorized as men both quickly and accurately. These pre-dissertation studies were instrumental in developing the hypotheses and theory for the following studies. I had two main hypotheses:

Hypothesis 1: Skin Tone Bias

A. **Skin tone bias in race categorization:** Darker skin tone will facilitate the categorization of Black men and women into their racial group compared to lighter skin.

B. Skin tone bias in gender categorization: Darker skin will facilitate the categorization of Black men as members of their gender and lighter skin will facilitate the categorization of Black women as members of their gender.

Hypothesis 2: Gendered Skin Tone Bias

A. Gendered skin tone bias in race categorization: Skin tone will moderate the effect of gender on racial categorization. Darker skin will facilitate the categorization of darker-skinned women (DW) vs. lighter-skinned women (LW), relative to darker-skinned men (DM) vs. lighter-skinned men (LM).

B. Gendered skin tone bias in gender categorization: Skin tone will moderate the impact of gender on gender categorization. Participants will be better at categorizing LW vs DW, relative to DM vs LM. In other words, skin tone will facilitate Black women's gender categorization to a greater extent than Black men.

Methods

Participants

Participants were recruited from the psychology course participant pool of a northeastern university in exchange for course credit. I calculated sample sizes using the PANGEA program (Westfall, 2016) for analyzing sample size and power for ANOVA designs with fixed and random factors, to have at least 80% power to detect a small-to-medium effect size and determined at least 100 participants were necessary. I recruited 107 participants for Study 1a: 60 participants were White, 28 Asian (East Asian and Southeast Asian), two participants were of Middle Eastern descent, eight participants were Black or African American (non-Hispanic), five participants identified as multiracial and/or other, three identified as Latino/Hispanic (regardless

of race), and one participant was Native American. The sample had 76 women and 31 men. I recruited 115 participants for Study 1b: 71 participants were White, 18 identified as multiracial, 18 participants were Asian (East Asian and Southeast Asian), one was of Middle Eastern descent, five were Black or African American (non-Hispanic), two were Latino/Hispanic (regardless of race). The sample had 78 women and 37 men. Participants in both samples completed the same procedure and the results were similar, so these studies are presented simultaneously.

Procedure

Participants were told via computerized directions that they would be categorizing individual Black and White faces by their race first or gender first, and which task came first was randomized. They were instructed to categorize the faces as quickly as possible by clicking equidistant buttons on a keyboard, while making as few mistakes as possible. After an eight-trial practice round, participants categorized 40 Black male and female faces (20 lighter-skinned and 20 darker-skinned) and 40 White male and female faces in 128 trials per categorization task, 256 trials in total. To increase generalizability and power, in Study 1b, I increased the number of faces participants saw to 160 (40 LW, LM, DW, DM). Participants categorized 480 trials of Black and White, male and female faces in the race categorization task and 240 trials of the gender categorization task. Participants then completed measures of racial attitudes and a demographic survey, after which they were thanked for their participation and given course credit.

Materials

I collected front-facing, neutral expression photos of Black and White men and women and cropped to a circle outlining the face. Most facial stimuli (21 faces ~50%) were collected

from the Chicago Face Database (Ma et al., 2015). I collected nine faces (23%) from the Multi-Racial Mega-Resolution database (Strohminger et al., 2016). Two photos were from the Lifespan database of adult facial stimuli (Minear & Park, 2004). I collected five additional photos from the NBA and WNBA 1998-2000 team rosters and chose these years so that the faces in question would be unlikely to be recognized by participants. The remaining faces (two) were taken from volunteers at two universities. To determine skin tone, I used the Massey & Martin Skin Color scale (Massey & Martin, 2003). Faces at or below a three were categorized as lighter-skinned and faces at or above a six were considered darker-skinned. This resulted in 10 LW, 10 LM, 10 DW, and 10 DM. To increase the number of stimuli for Study 1b, I used facial morphing and composite software (DeBruine, 2018) to make 10 additional faces for each category, made from combinations of the first 10 faces. This resulted in 80 faces total, 20 facial stimuli per category (light/dark, male/female).

Results

For both studies I analyzed both the reaction time and error rate using linear mixed effects models (LMEMs). As reported in Bart-Plange (2019), I included fixed effects for target skin tone, target gender, and their interaction, as well as a by-subject random intercept and random effects slopes for both effects and their interaction. I interpreted the models using the anova() function in R (Fox & Weisberg, 2019). This procedure uses Satterthwaite approximation method of calculating degrees of freedom and F statistic, and accounts for the nonindependence of participants' multiple observations in the data[1]. LMEMs also often estimate denominator degrees of freedom with decimals. In the reaction time analyses, I only analyzed correct responses. Latencies under 300ms or over 3,000ms were re-coded to 300ms and 3,000ms respectively (Greenwald et al., 2003). I transformed the latency (ms) to reaction time (s) to

[1]There is no consensus among statisticians on the best standard of computing degrees of freedom (Baayen et al., 2008). However, the Satterthwaite approximation is a commonly used and is the default approximation in SAS, SPSS, and the anova() function in R.

obtain a normal distribution (Whelan, 2010). The subsequent Box-Cox analysis showed a $\lambda = 1$, indicating that no further transformations of the data were necessary.

Race Categorization Reaction Time and Accuracy

Study 1a

Accuracy Analysis. For rate of categorization errors, there was a main effect of skin tone $F(1,113.57) = 119.25$, $p < .001$, such that participants were more accurate in categorizing darker-skinned faces compared to lighter-skinned faces, supporting H1-A[2]. There were no effects of gender or an interaction, therefore we found no evidence of gendered skin tone bias in race categorization accuracy, contrary to H2-A, $F(1,245.16) = 0.59$, $p = .444$; $F(1,220.51) = 0.24$, $p = .624$. See Figure 1.

Reaction Time Analysis. Results revealed only a main effect of skin tone, consistent with H1-A; participants were quicker to categorize darker-skinned target faces compared to lighter-skinned targets, $F(1,107.07) = 851.13$, $p < .001$. I found no gender effects or interaction, contrary to H2-A, meaning that there was no evidence of gendered colorism in the speed of categorization analysis, $F(1,657.20) = 2.52$, $p = .113$; $F(1,385.36) = 0.0074$, $p = .932$. See Figure 2.

Study 1b

Accuracy Analysis. Study 1b participants made less errors in categorizing darker-skinned faces as Black than lighter-skinned faces, supporting H1-A, $F(1,146.94) = 322.089$, $p < .001$. There was an effect of gender, $F(1,163.81) = 13.109$, $p < .001$, and an interaction of race and gender, such that skin tone had a larger impact on the incorrect categorization of female faces compared to male faces $F(1,178.91) = 14.53$, $p < .001$, supporting H2-A. For race categorization, this was the only evidence of gendered skin-tone. Holm-adjusted pairwise

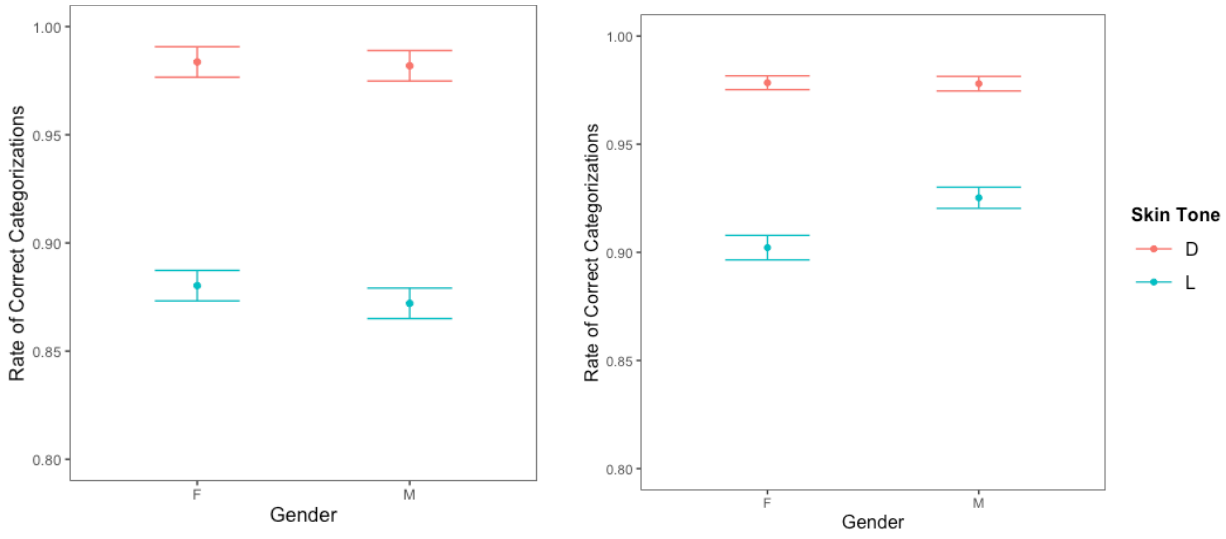
[2]This model failed to converge with 1 negative eigenvalue: -3.9e-02. I removed the random intercepts of skin tone and gender, after which the model converged and the results remained the same.

contrasts showed that darker-skinned vs. lighter-skinned Black faces were categorized accurately more often, DF-LF: $t(136) = 14.85$, $p < .001$; DM-LM: $t(205) = 12.29$, $p < .001$, and the skin tone effect was stronger for Black women, $t(179) = 3.81$, $p = .003$. In other words, participants had the greatest difficulty with accuracy of categorizing lighter-skinned Black people when they were women. See Figure 1.

Reaction Time Analysis. I found a main effect of skin tone, supporting H1-A, and a main effect of gender, such that participants were faster to categorize darker-skinned people relative to lighter-skinned people and Black men relative to Black women, $F(1,112.67) = 900.71$, $p < .001$; $F(1,3006.17) = 35.67$, $p < .001$. There was no interaction showing evidence of a gendered skin tone bias to support H2-A, $F(1,116.39) = 1.85$, $p = .176$. See Figure 2.

Figure 1

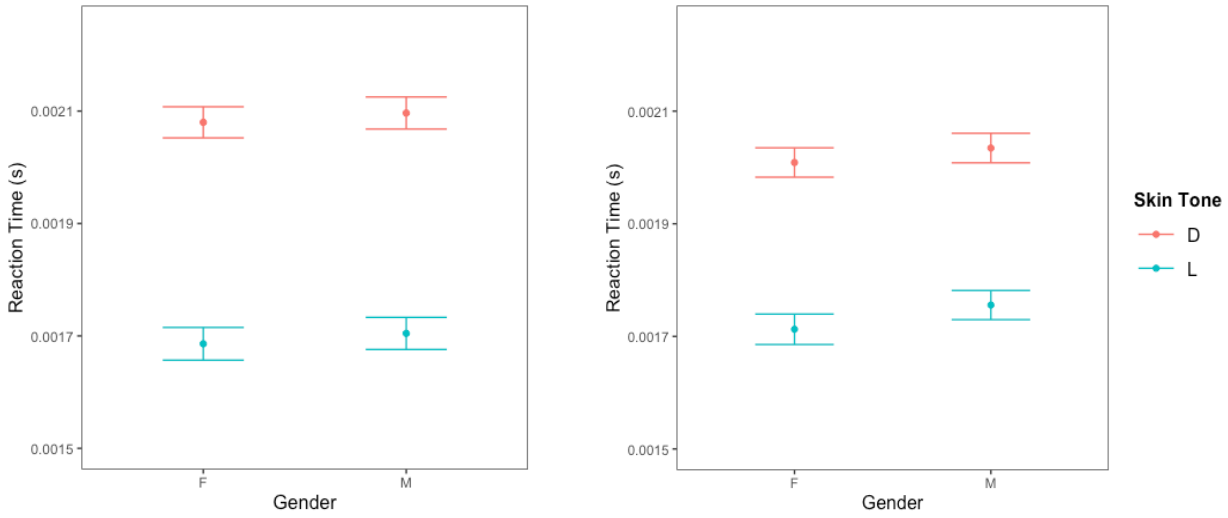
Study 1a (left) and 1b (right) Mean Racial Categorization Accuracy of Black Faces



Note. Error bars represent the standard error of the mean.

Figure 2

Study 1a (left) and 1b (right) Reaction Time of Racial Categorization of Black Faces



Note. Error bars represent the standard error of the mean.

Gender Categorization Reaction Time and Accuracy

Study 1a

Accuracy Analysis. Supporting H1-B and H2-B, analyses revealed main effects of skin tone and gender, $F(1,144.22) = 54.94$, $p < .001$; $F(1,146.52) = 75.16$, $p < .001$, and an interaction, $F(1,276.15) = 150.23$, $p < .001$. Male faces were categorized more accurately than female faces, and darker-skinned faces were categorized more accurately than lighter-skinned faces. Subsequent holm-adjusted pairwise contrasts showed that dark skin facilitated gender categorization of Black men, $t(277) = 3.13$, $p = .002$, but hindered gender categorization for Black women $t(149) = -13.11$, $p < .001$, and that the extent to which skin tone influenced these errors was larger for Black women than men, $t(276) = -12.26$, $p < .001$. See Figure 3.

Reaction Time Analysis. Supporting H1-B, analyses revealed main effects of skin tone and gender, $F(1,197.53) = 6.17$, $p = .014$; $F(1,105.19) = 90.94$, $p < .001$, qualified by an interaction, $F(1,308.40) = 100.87$, $p < .001$. Subsequent holm-adjusted pairwise contrasts

showed that participants categorized darker- vs. lighter-skinned Black men as men more, $t(108) = 8.72$, $p < .001$, and darker- vs. lighter-skinned Black women as women more slowly, $t(2674) = -5.36$, $p < .001$. The former difference was most pronounced, which was in the opposite direction predicted in H2-B, $t(3080) = -10.043$, $p < .001$. See Figure 4.

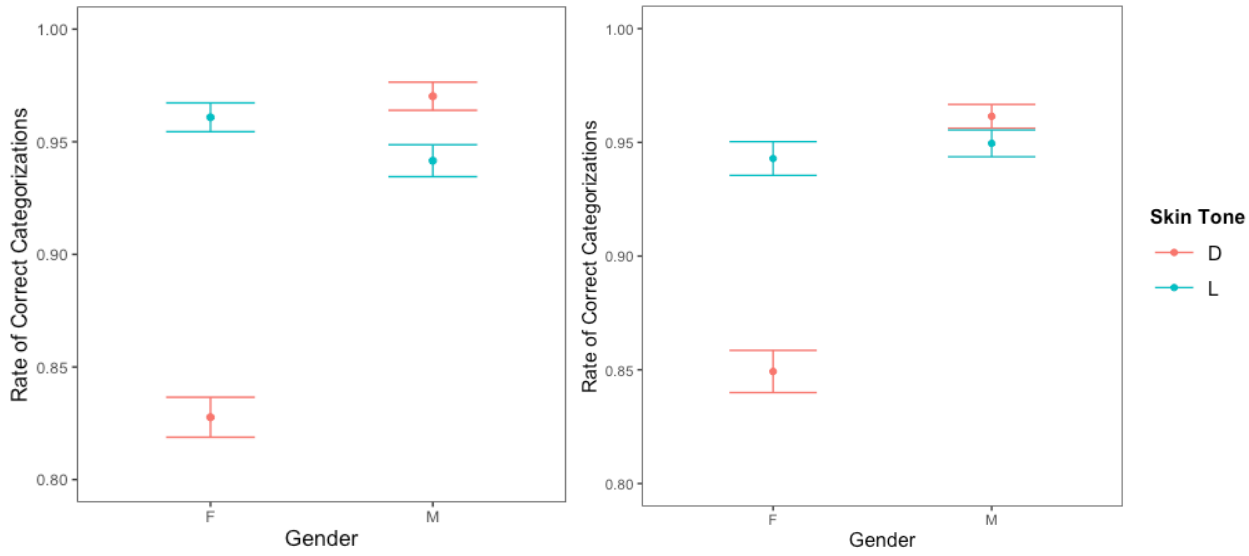
Study 1b

Accuracy Analysis. Study 1b results replicated Study 1a main effects of skin tone and gender, supporting H1-B and H2-B, $F(1,132.11) = 57.70$, $p < .001$; $F(1,114.92) = 84.37$, $p < .001$, and the interaction driving these effects, supporting H2; $F(1,194.73) = 122.91$, $p < .001$. This again showed evidence of gendered colorism. Holm-adjusted pairwise contrasts confirmed that darker skin hindered the categorization of Black women; $t(115) = -11.68$, $p < .001$, and marginally facilitated the categorization of Black men, $t(584) = 1.90$, $p = .057$, and that this difference was most pronounced for Black women, $t(195) = -11.087$, $p < .001$. See Figure 3.

Reaction Time Analysis. Results also replicated Study 1a and supported H1-B; results showed main effects of skin tone and gender, $F(1,111.81) = 58.35$, $p < .001$; $F(1,113.27) = 90.96$, $p < .001$, and the same interaction, $F(1,110.57) = 94.30$, $p < .001$, such that dark skin facilitated the categorization of Black men but hindered the categorization of Black women. Subsequent holm-adjusted pairwise contrasts showed that darker-skinned Black women were categorized marginally slower than lighter-skinned Black women, $t(108) = -2.14$, $p = .070$, and darker-skinned Black men were categorized faster than lighter-skinned Black men $t(116) = 11.95$, $p < .001$. Skin tone effects were again more pronounced for Black men, $t(112) = -9.76$, $p < .001$. See Figure 4.

Figure 3

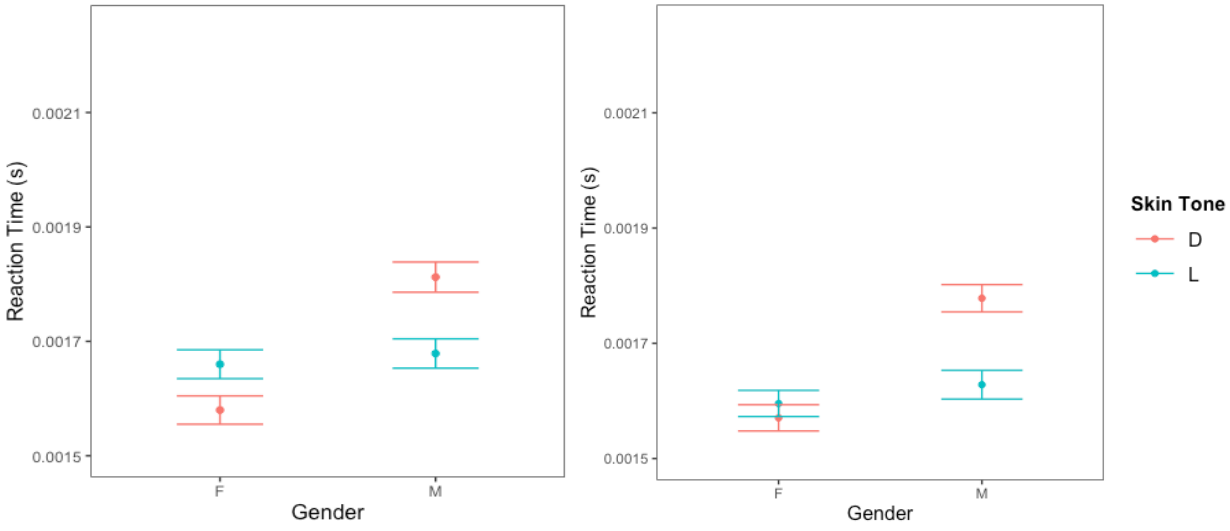
Study 1a (left) and 1b (right) Mean Accuracy of Gender Categorization of Black Faces



Note. Error bars represent the standard error of the mean.

Figure 4

Study 1a (left) and 1b (right) Reaction Time of Gender Categorization of Black Faces



Note. Error bars represent the standard error of the mean.

Discussion

In Studies 1a and b, consistent with H1-A, darker skin tone consistently facilitated the categorization of men and women into the category “Black”, and the categorization of men as “Male”. The gender effect and interaction in the racial categorization task was inconsistent—I predicted that consistent with previous work (Goff, Thomas, & Jackson, 2008; Goff, Steele, & Davies, 2008), participants would be better at categorizing Black men compared to Black women, darker-skinned people compared to lighter-skinned people (H1-A), and that the effect of skin tone would be larger for Black women (H2-A). However, the gender effect and interaction were only present in one study and thus did not provide consistent evidence for a gendered skin tone bias. Skin tone did consistently matter for successful racial categorization; participants were quicker and more accurate to categorize darker-skinned vs. lighter-skinned faces. Because race is a gendered construct, I expected that a gendered skin tone bias would occur in the racial categorization task as well, predicting that race would similarly be a gendered domain. It may be that because the racial categorization was a reaction time task and these categories were explicitly displayed, the most race stereotypic feature, skin tone, was the most salient and overpowering factor in race categorization.

Gender categorization accuracy and speed results were consistent with H1-B in both Studies 1a and b: participants categorized Black men as men faster and more accurately, particularly for men with darker skin. Participants categorized Black women as women slower overall, especially Black women with darker skin. Consistent with H2-B, skin tone disproportionately impacted the gender accuracy of Black women and hindered participant’s speed in gender categorization. The extent to which skin tone influenced the speed of categorization was larger for Black men than Black women, differing from the accuracy data. Some reasons for this divergence from accuracy results is that because participants were so much

slower to categorize Black women overall, there is less room for major differences in speed for categorizing lighter and darker-skinned Black women, but more so for Black men, who are categorized as male faster overall (i.e., a floor effect).

These studies consistently evidence the strong associations of "Blackness" with "maleness" on a categorical level, and therefore why Black women's perceived non-prototypicality is of greater consequence for them when gender is salient. For Black men, although lighter skin led to slower gender categorizations, accuracy remained high (Study 1a: 94%, Study 1b: 95%). For Black women, darker skin tone led to both slower and substantially less accurate gender categorization compared to all other groups (Study 1a: 83%, Study 1b: 85%). In this way, when the domain of gender was salient, the effect of skin tone on accuracy was more consequential for Black women. In other words, skin tone consistently had the largest effect on darker-skinned Black women's perceived "woman-ness". However, in the racial categorization task, when race instead of gender was salient, the results consistently showed a skin tone bias; participants were much more successful associating darker skin with "Blackness".

These studies provided insight into how skin tone influences categorical processing, laying the groundwork for Study 2. Although these studies provided insight into how skin tone complicates how quickly and accurately Black men and women are understood as a member of their race or gender ingroups, they did not speak to whether these features influence social judgments. In Study 2, I examined the effect of gendered domains on prototypicality saliency and examine how skin tone biases evaluations of Black men and women.

Study 2: Prototypicality Salience in Gendered Domains

In Study 2, I further examined how skin tone shapes social cognition; specifically, I investigated whether people naturally use skin tone in person perception. After being primed

with a gender-salient video or a neutral video, participants rated how similar different White and Black faces looked. I used multidimensional scaling techniques to investigate how a gendered context vs. neutral control may change how similar participants view lighter- and darker-skinned Black women to each other, to White women, and to Black men. This task was designed to examine how skin tone would be used to determine similarity within a racial group and across racial groups in a neutral and a gender-salient context. Because skin tone intersects with race and gender, I predicted that skin tone will be used to determine similarity within a racial group and across race. I also predicted that darker-skinned Black men and women and White men and women will be seen as the least similar to each other.

I also directly tested the assumption that skin tone biases perceptions of race and gender prototypicality. Here, participants rated White and Black faces varying in skin tone on how stereotypical they looked of their race and gender. Because "Blackness" is associated with "maleness", Black women have stereotypes attributed to them which are distinct from stereotypes of "Black people", which tend to apply to Black men, and of "women", which tend to apply to White women (Goff, Thomas, & Jackson, 2008; L. K. Johnson et al., 2012). Race and gender are basic categories of person perception (Fiske & Neuberg, 1990). As skin tone is the primary indicator of race, it influences categorizations of Black men and women's racial and gender identities (as evidenced in Study 1) and should also influence how prototypical Black men and women are evaluated as members of their race and gender. Therefore, because Black women are seen as non-prototypical group members in ways that Black men are not, skin tone should bias perceptions of their gender prototypicality to a greater extent. Black women (especially DW) will be seen as less prototypical of women than White women. Because Task 1 was relatively long—participants completed similarity ratings for 66 pairs of faces, I did not

expect the gender-salient video to carry over and affect participants' ratings of prototypicality. Thus, I tested the following hypotheses in Study 2:

Hypotheses

H1: Skin Tone Bias

Skin tone will bias evaluations of Black men and women, such that:

A. Skin tone will matter in ratings of similarity, such that participants will find faces of similar skin tone as more similar.

B. Participants will find darker-skinned faces of Black men and women as more prototypical of their race compared to lighter-skinned faces.

C. Participants will find DM as more prototypical of men compared to LM and LW as more prototypical of women than DW.

H2: Gendered Colorism

Skin tone will bias evaluations of Black women to a greater extent than Black men when gender is salient such that:

A. When gender is made salient, DW and LW will be rated as less similar to each other compared to the neutral control because lighter-skinned Black women (LW) will be rated as more similar to White women (WW) while darker-skinned women (DW) will be rated as more similar to darker-skinned Black men (DM).

B. In the gendered domain (prototypicality of gender), the extent to which skin tone biases perceived prototypicality of gender will be larger for Black women than for Black men.

Methods

Participants

I recruited 463 participants from the participant pool of a large northeastern American university. I conducted a sensitivity analysis using the G*Power program (Faul et al., 2009), and determined I had enough power to detect a small effect (d = .18). Participants were 134 men, 326 women, and 3 non-binary individuals. The final sample had 295 White participants, 21 Black, 106 participants were Asian (East, Southeast, or Pacific Islander), and 4 were of Middle Eastern descent. Thirty-two participants identified as Latino/a, regardless of race, and participants were on average 19 years old.

Stimuli

The stimuli were front-facing, neutral expression photos of Black and White men and women. I matched each group's photos for average age and skin tone per skin tone and gender level (LW, LM, DW, DM) from the Chicago Face Database (Ma et al., 2015).

Procedure

After consenting to participate in the study, participants were instructed that they will be watching a video. I directed participants to pay close attention, as they would be quizzed on what they watched directly after. Participants watched either a video meant to prime gender salience (three minute and 45 second video about popular women's hairstyles from 1920 - 1960; Allure, 2020) or a neutral video (three minute and 30 second video about caterpillars turning into butterflies; Science Insider, 2019). After watching the video, participants answered a series of questions about the video they watched, which served as an attention check and reinforcement of gender saliency in the gendered salience condition. Directly after this, participants completed two tasks:

Task 1: Similarity Task

Participants were presented with pairs of faces and asked to rate how similar each pair is on a Likert scale (1 = *Not at all similar* to 7 = *Very similar*). Each photo was of a lighter- or darker-skinned Black woman or man, or White man or woman. Participants rated the similarity of 12 faces in total (2 of LW, DW, DM, LM, WW, WM), in 66 total pairings.

Task 2: Prototypicality Rating Task

Participants were then told that they would view the same faces as the previous task, one at a time, and answer questions about their impressions of them. Participants viewed all 12 faces from the previous similarity task and answered, "How stereotypically Black/White does this person look?" and "How stereotypically female/male does this person look?" (1 = *Not at all* to 6 = *Very much*). Participants then completed measures of racial attitudes and a demographic survey, after which they were thanked for their participation and given course credit.

Results

Task 1: Similarity Task

To determine whether participants showed differences in similarity ratings depending on condition (gender salient vs. gender neutral), I created a linear mixed-effects model with domain and comparison as fixed effects, a by-subject random intercept, and random effects slopes for both effects and their interaction. I interpreted the model using the anova() function in R, which approximates degrees of freedom with the Satterthwaite approximation. Contrary to prediction H2-A, results showed no effect of condition, $F(1,472.2) = 0.68$, $p = .41$. There were, however, differences in how similar participants rated different pairs of faces, $F(20,30049) = 552.77$, $p < .001$. Post-hoc holm-adjusted pairwise contrasts confirmed that LM and LW were seen as equally dissimilar to WM and WW respectively, but DW and WW were seen as less similar to each other than DM and WM. Lastly, LW were considered more similar to WW than DW to

WW, and DW were considered more similar to DM than LW to DM. See Table 1. This provided evidence for gendered colorism but was not consistent with the differences by condition proposed in H2-A. See Figure 5. As there was no effect of condition, I presented the results for the MDS visual representation and similarity ratings combined across conditions. See Table 2.

Table 1

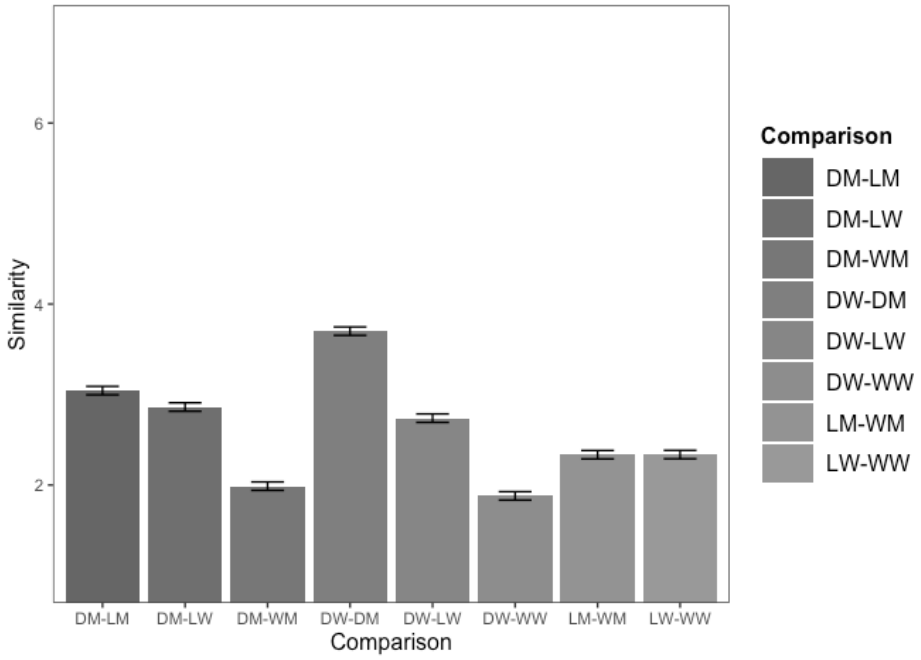
Holm-Adjusted Pairwise Contrasts of Similarity Pair Ratings

Contrast	SE	df	*t* ratio	*p* value
(LM-WM)-(LW-WW)	0.035	14335	-0.038	.970
(DM-WM)-(DW-WW)	0.035	14335	51.78	< .001**
(DW-WW)-(LW-WW)	0.035	14335	-12.99	< .001**
(LW-DM)-(DW-DM)	0.035	14335	-23.86	< .001**

Note. $p < .05$*, $p < .001$**

Figure 5

Similarity Pairings Averaged Across Condition



Note. Error bars represent the standard error of the mean.

Table 2

Mean Similarity Pairings Between Black and White, Male and Female Faces

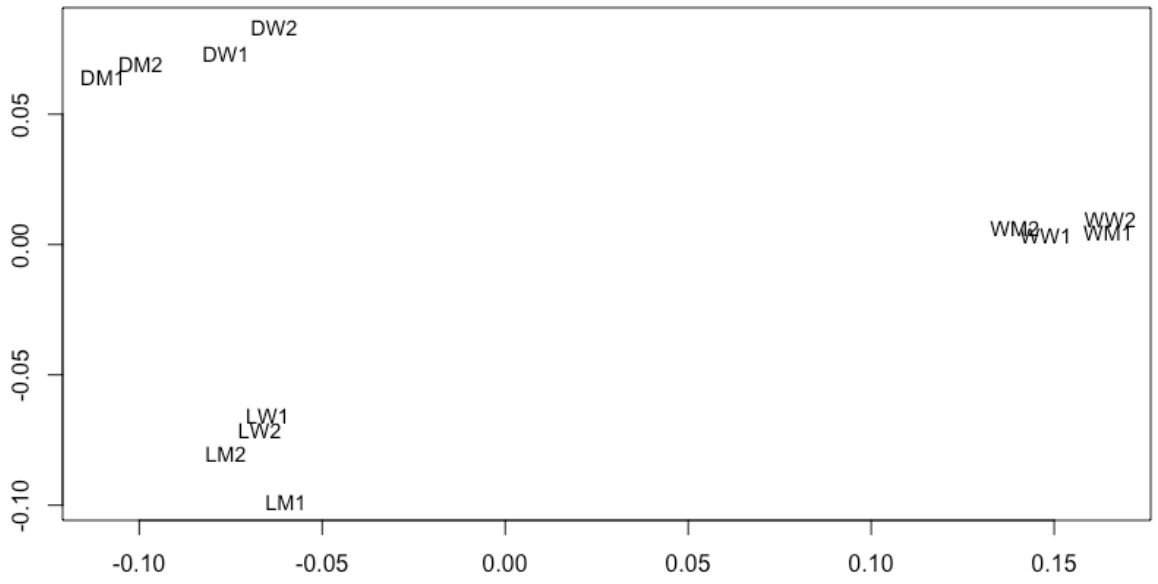
Comparison	Mean	SE	95% CI lower	95% CI upper
LM-LM	4.29	0.06	4.17	4.42
WW-WW	3.90	0.06	3.77	4.02
LW-LW	3.81	0.06	3.69	3.94
WM-WM	3.78	0.06	3.66	3.91
DW-DM	3.70	0.05	3.61	3.79
DM-DM	3.47	0.06	3.34	3.59
LW-LM	3.42	0.05	3.33	3.51
WW-WM	3.29	0.05	3.20	3.38
DW-DW	3.12	0.06	3.00	3.25
DM-LM	3.04	0.05	2.95	3.13
DM-LW	2.86	0.05	2.77	2.95
DW-LW	2.74	0.05	2.65	2.83
DW-LM	2.55	0.05	2.46	2.64
LW-WW	2.34	0.05	2.25	2.43
LM-WM	2.34	0.05	2.25	2.43
LM-WW	2.26	0.05	2.17	2.35
LW-WM	2.14	0.05	2.05	2.23
DM-WM	1.99	0.05	1.90	2.08
DW-WM	1.95	0.05	1.86	2.04
DW-WW	1.88	0.05	1.79	1.97
DM-WW	1.75	0.05	1.66	1.84

MDS Dimensionality and Space

I analyzed the MDS results using the MASS and Vegan packages and algorithms in RStudio (Venables & Ripley, 2002; Oksanen et al., 2020). The matrix stress analysis was nearly zero (8.86173e-05), which indicated that there may have been several other dimension solutions or that more observations were necessary to evaluate the data. I then used metric scaling, or Classical Multidimensional Scaling (CMDS) to create the visual distance plot of the data. See Figure 6. The two-dimensional CMDS solution did not provide evidence for H2-A as there was no significant difference between the gendered and neutral condition, but rather showed that supporting H1-A, participants rate the similarity of others clustering around race and skin tone. Within these groups, participants generally clustered around gender.

Figure 6

Two-dimensional Classical Multidimensional Scaling Visualization



Note. The x- and y-axes represent the primary and secondary dimensions, respectively.

Perceived Prototypicality

To analyze the perceived prototypicality ratings of race and gender, I constructed a linear mixed-effects model including fixed effects for target skin tone, target gender, condition, and their three-way interaction, as well as a by-subject random intercept and random effects slopes for the interaction between skin tone and gender for each model. I interpreted the model using the Anova() function in R, which approximates degrees of freedom and the F statistic using Kenward-Roger's method of approximation, a modification of the Satterthwaite approximation.[3] Participants had watched the gender salient or gender neutral video and completed the similarity task prior to task two, and as expected, participants ratings did not differ by condition, $F(1,460.85) = .097$, $p = .755$.

Race Prototypicality

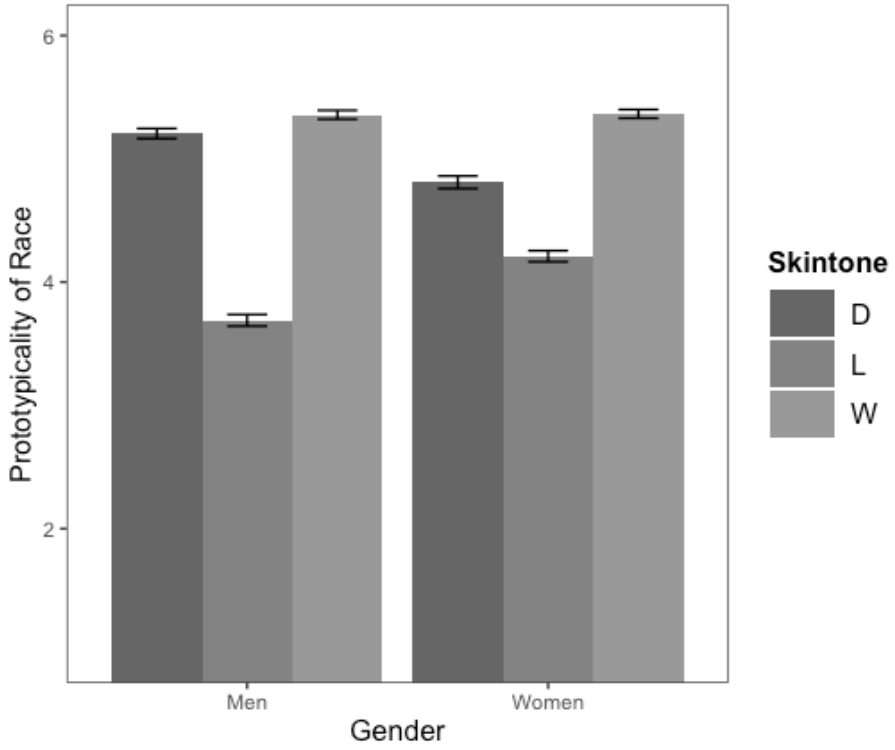
For the race prototypicality mixed model analysis, supporting H1-B, there were main effects of skin tone and gender, $F(2,460.16) = 624.85$, $p < .001$; $F(1,460.13) = 106.98$, $p < .001$, and an interaction between skin tone and gender, $F(2,460.21) = 137.60$, $p < .001$, such that the effect of skin tone on race prototypicality was more pronounced for Black men than Black women. White men and women were found as more prototypical of their group than Black men and women regardless of skin tone, WM-DM: $t(462) = 4.10$, $p = .001$; WM-LM; $t(462) = 33.05$, $p < .001$; WW-DW: $t(462) = 12.06$, $p < .001$; WW-LW: $t(462) = 25.01$, $p < .001$. As expected and supporting H1-B, for Black targets, DM were seen as the most prototypical of their racial group overall, and darker-skinned men and women were seen as more prototypical of their race compared to lighter-skinned Black men and women, DM-LM: $t(462) = 32.40$, $p < .001$; DW-LW: $t(462) = 12.03$, $p < .001$. In Studies 1a & b, evidence for a gendered skin tone bias for Black women was mixed, thus I did not have a clear expectation of whether there would be a gendered

[3]The Kenward-Roger degrees of freedom approximation will rescale the F ratio and degrees of freedom from the Satterthwaite approximation in such a way that results in a more appropriate F distribution (Judd et al., 2012). This method takes considerably more computer power to perform (Kuznetsova et al., 2017) and is not always possible to compute for larger models with substantially more observations.

skin tone bias when race was salient. However, skin tone mattered more for the racial evaluation of Black men compared to Black women. See Figure 7.

Figure 7

Perceived Racial Prototypicality by Skin Tone and Gender of Targets



Note. Error bars represent the standard error of the mean.

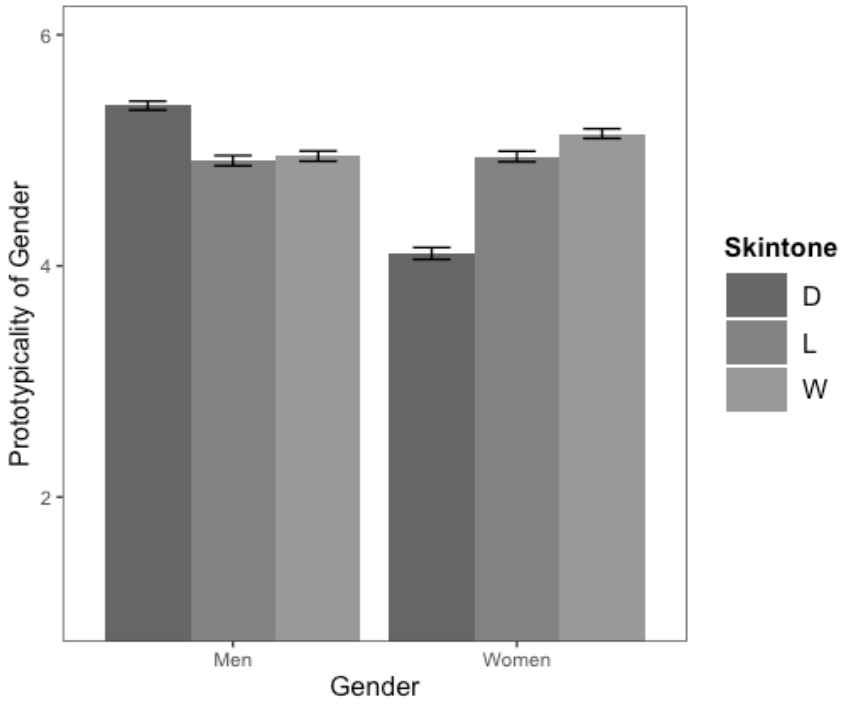
Gender Prototypicality

For the gender prototypicality mixed model analysis, supporting H1-C and H2-B, there were main effects of skin tone and gender, $F(2,460.81) = 73.83$, $p < .001$; $F(1,461.81) = 629.02$, $p < .001$, and an interaction between skin tone and gender, $F(2,461.74) = 277.05$, $p < .001$. Further investigation of the interaction revealed that darker skin heightened the extent to which participants viewed Black men as prototypical of men, DM-LM: $t(462) = 11.15$, $p < .001$; DM-WM: $t(462) = 9.57$, $p < .001$; LM-WM: $t(462) = -0.87$, $p = .954$, and hindered the extent to which participants viewed Black women as prototypical of women, DW-LW: $t(462) = -17.98$, p

< .001; DW-WW: $t(462) = -21.27$, $p < .001$; LW-WW: $t(462) = -4.67$, $p < .001$. As expected and supporting H2-B, the effect of skin tone significantly mattered more for perceived prototypicality of women's gender, $t(461) = 16.57$, $p < .001$ and DW were seen as the least prototypical of women. See Figure 8.

Figure 8

Perceived Gender Prototypicality by Skin Tone and Gender of Targets



Note. Error bars represent the standard error of the mean.

Discussion

Study 2 examined the role of gender saliency in colorism in two tasks. Task 1 results supported H1-A. Participants clearly used skin tone to make their similarity judgments. The MDS visualization revealed that participants mainly judged similarity on the dimensions of race and skin tone, and then within these groupings, gender. Task 1 results also partially supported H2-A. LW were considered more similar to WW than DW to WW, and DW were considered

more similar to DM than LW to DM, although this was not pronounced in the gender salient condition, contrary to prediction. Moreover, gender, and thus, gender non-prototypicality, seemed to be salient to participants as they rated the similarity between groups. Lighter-skinned Black men and women were seen as equally dissimilar to White men and women, but darker-skinned Black women and White women were seen as more dissimilar to each other than darker-skinned Black men and White men. Thus, skin tone mattered more for evaluating the similarity between DW and WW. In other words, skin tone mattered more for evaluating Black women's similarity to those deemed most prototypical of women. This is consistent with gendered colorism and suggests that gender may need to be inherent to the task or question in order to elicit gendered colorism.

Task 2 ratings provided support for H1-B, H1-C, and H2-B. In the racial prototypicality rating task, darker-skinned Black men and women were seen as more prototypical of their race compared to lighter-skinned Black men and women. This was consistent with H1-B. Participants also used skin tone as a marker of racial prototypicality to a greater extent for Black men compared to Black women: DM were seen as highly prototypical of Black compared to LM and Black women overall. However, in the gender prototypicality rating task, skin tone mattered much more for the perceived prototypicality of Black women than Black men. The skin tone bias observed for Black men in the racial prototypicality task did not translate to the same erasure of their gender, or gendered colorism, observed for Black women. This was quite consistent with H2-B, that when gender is salient, skin tone will matter more for evaluations of Black women. Both lighter- and darker-skinned Black women were seen as less prototypical of women than White women, also supporting my hypotheses and consistent with prior literature (L. K. Johnson et al., 2012).

Overall, these results mirrored the race and gender categorization data of Studies 1a and b and supported the hypothesis that Black women are perceived as dually non-prototypical of their race and gender, providing evidence that when gender is made salient, skin-tone matters more for both the categorization and evaluation of Black women.

Study 3: Gendered Colorism in Pain Perception

In Study 3, I moved from perceptual biases in categorization to evaluative judgments from Studies 1 and 2 to pain perception. Here, I examined whether people assume that darker-skinned Black men and women feel less pain than lighter-skinned Black men and women, and whether colorism is more pronounced for Black women in gendered painful scenarios, meaning that the scenario is thought of to be more typical of women (e.g., I burn myself on a curling iron") or more typical of men (e.g., I nick my chin while shaving). In this study, participants rated how much pain Black men and women would feel for pre-tested gendered scenarios and neutral scenarios.

I also examined factors that predict racial pain bias: prototypicality and assumed life hardship. Previous work has shown that racial prototypicality exacerbates racial bias in pain perception: Black people with more Afrocentric features are perceived to feel less pain (Drain et al., 2020). Because race and gender are interrelated constructs, I expected that prototypicality of gender should 1) also influence pain ratings, and 2) influence pain ratings to a greater extent for Black women (vs. men) if gender salience is of greater consequence for them. Previous work has also shown that perceptions of life hardship mediate racial bias in pain perception. That is, people believe Black people feel less pain than do White people to the extent that they believe life hardship "toughens" people (K. M. Hoffman & Trawalter, 2016). Because Black Americans experience racism, lay beliefs about life hardship and toughness leads to assumptions that Black

people will then feel less pain than White Americans. I predicted that people who are more prototypical of "Black" (i.e., darker-skinned Black people), will be assumed to have faced more life hardship, and therefore feel less pain. I also expected that because darker-skinned Black women are perceived as less prototypical of women, this would also predict beliefs of greater life hardship and feeling less pain, particularly in scenarios where gender is salient. This study presented multiple ways to test the two main predictions of finding evidence of colorism and gendered colorism. I again tested the two main predictions from the proposed theory:

H1: Skin Tone Bias

Skin tone will bias evaluations of Black men and women, such that:

A. Participants, on average, will rate the pain of darker-skinned Black people as less severe than the pain of lighter-skinned Black people.

B. Participants will rate darker-skinned Black people as more prototypical of their racial group compared to lighter-skinned Black people.

C. Participants will rate DM as more prototypical of men vs. LM, and LW as more prototypical of their gender vs. DW.

D. Participants will assume darker-skinned Black people relative to lighter-skinned Black people have experienced more life hardship.

E. Race prototypicality and/or assumed hardship will mediate the relationship between skin tone and pain ratings.

H2: Gendered Colorism

Skin tone will bias evaluations of Black women to a greater extent than Black men when gender is salient, such that:

A. In the gender-salient pain scenarios, the discrepancy between the ratings of pain for lighter vs. darker-skinned Black people will be most pronounced for Black women.

B. Skin tone will be used to evaluate gender prototypicality to a greater extent for Black women.

C. For gender-salient pain scenarios, prototypicality of gender will mediate the relationship between skin tone and pain ratings to a greater extent for Black women.

Method

Participants

I recruited 674 participants from two northeastern American universities participant pools for class credit. I calculated sample sizes using the PANGEA program (Westfall, 2016) for analyzing sample size and power for ANOVA designs with fixed and random factors, to have at least 90% power to detect a small-to-medium effect size. I accounted for the random factors of participant's multiple observations of individual pain scenario items and determined 650 participants was sufficient to detect a three-way interaction between skin tone, gender, and domain. 204 participants were men, 462 were women, 13 participants were nonbinary, one identified as genderqueer, and one did not provide their gender identity. The sample was 351 White, one identified as Native American, 84 participants were Black or African American (not Hispanic), 48 Hispanic and/or Latino (regardless of race), 117 participants were Asian (East Asian, Southeast Asian, and Pacific Islander), 69 identified as multiracial, and four declined to list their race or ethnicity.

Procedure

After consenting to participate, participants rated the amount of pain they would feel in a series of 16 every-day scenarios on a four-point scale (1= *Not painful* to 4 = *Extremely painful*; α

= .82). Eight items were non-gender salient painful situations and eight were gender salient painful situations. Examples of non-gender salient painful situations were adapted from K. M. Hoffman & Trawalter (2016), including scenarios such as "I stub my toe on a chair leg" and "I disinfect a sore" and gender salient painful situation examples included statements such as "I burn myself on a curling iron" and "I nick myself on the chin shaving" (α = .81). Gendered pain items were pre-tested for laypeople's beliefs on how likely each scenario is to happen to a man or a woman and matched for pain intensity with the neutral items (See Table 2). Participants then viewed a picture of a lighter- or darker-skinned Black man or woman and completed the same survey asking them to rate how much pain they believed the person presented would feel in each scenario. Participants made these ratings for four target persons total, two men and two women, one lighter- and darker-skinned of both genders. Participants then rated the same faces on the prototypicality of their race and gender by answering "How stereotypically Black does [name] look?" and "How stereotypically female/male does [name] look?" on a six-point scale (1 = *Not at all stereotypical* to 6 = *Very stereotypical*), and "How light- or dark-skinned is [name] for the average African American?". Lastly, participants answered a series of four questions about how much hardship and adversity they believed each target person had experienced in life (K. M. Hoffman & Trawalter, 2016), which were, "How privileged do you think [name] is?", "How hard do you think [name]'s life has been?", "How lucky do you think [name] has been?", "How much adversity do you think [name] has overcome?" on a five-point scale (1 = *No adversity at all* to 5 = *Extreme adversity*). These questions were averaged into a composite score of assumed hardship, α = .70. Participants then completed individual difference measures and demographic information, after which they were thanked for their participation and debriefed.

Gender-Salient Scenarios

112 raters recruited from Amazon Mechanical Turk (MTurk; Crowston, 2012) read 18 painful scenarios and indicated whether the scenario was more typical of men or women. Participants also rated and how painful they believed these scenarios would be along with the 18 original painful scenario items from K. M. Hoffman and colleagues (2016). A final list of eight gendered items were matched for pain intensity with the neutral pain items, four of which rated as typical of men and four of which were rated as typical of women. See Table 2.

Table 2

Pre-Tested Pain Ratings for Neutral and Gendered Pain Items

Neutral	Rating	Gendered	Rating
I get an injection in the arm.	2	I nick myself on the chin shaving.^{M}	1.99
I cut myself with a sheet of paper.	2.31	I collide with a friend while trying to catch a football.^{M}	2.29
I stub my toe on a chair leg.	3.07	I burn myself on a curling iron.^{W}	3.02
I get my fingers caught in the car door.	3.48	I drop a barbell on my toe at the gym.^{W}	3.50
I knock my head on the corner of a piece of furniture.	3.14	I burn myself on the oven grate while baking.^{W}	3.14
I get hit on the head by a stray Frisbee.	2.32	I prick my finger on a sewing needle.^{W}	2.27
I get sunburned on my face.	2.63	I get bitten on the hand while playing with a dog.^{M}	2.75
I get a speck of dust in my eye.	1.95	My earring gets stuck in my sweater. as I try to pull it off.^{W}	2.04

Note. The subscript ^{M} denotes gendered scenarios pre-tested to be associated with men, while the subscript ^{W} denotes gendered scenarios pre-tested to be associated with women.

Results

I analyzed the pain perception data using linear mixed-effects models. Each model had fixed effects for skin tone, gender, and scenario type, and participants and individual pain scenario items as random effects with a by-subject and by pain scenario item random intercept. I also controlled for self-pain ratings in these models. I again tested these models using the Satterthwaite approximation method of calculating degrees of freedom which accounts for the nonindependence of participants' multiple observations in the data. The prototypicality and hardship models included fixed effects for skin tone and gender, and participants as a random effect with a by-subject random intercept.

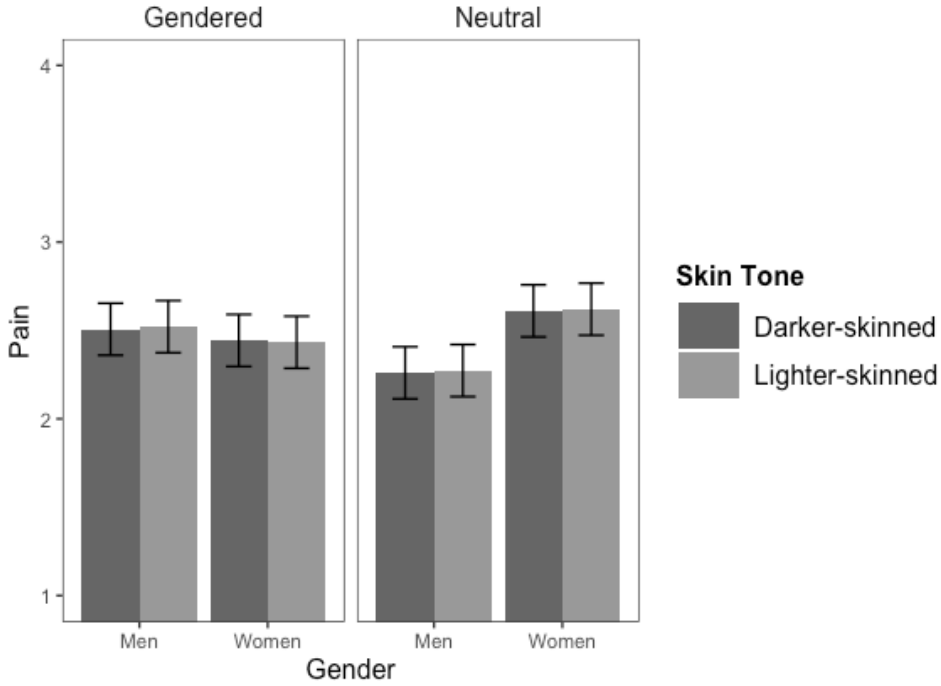
Pain Perception

Contrary to my expectations and H1-A, there were no effects of skin tone, $F(1,42882) = 0.93$, $p = .336$. Results revealed a main effect of gender, $F(1,42882) = 398.63$, $p < .001$, and an interaction between gender and pain domain, $F(1,42882) = 942.50$, $p < .001$. Subsequent Holm-adjusted pairwise comparisons showed that in gender neutral domains, participants gave higher pain ratings to Black women than to Black men's, $t(42886) = 35.67$, $p < .001$, and in gendered domains, participants gave lower pain ratings to Black women than to Black men, $t(42886) = -7.75$, $p < .001$[4]. A gender gap was present in both domains, however, H2-A predicting gendered skin tone bias in pain perception was not supported. See Figure 9.

[4]This interaction was largely driven by the feminine items. I examined gender differences on pain items that were masculine (e.g., I drop a barbell on my toe at the gym) vs. feminine (e.g., I prick my finger on a sewing needle). Results showed that among gender-salient pain scenarios, there was a main effect of gender of target, $F(1,21098) = 62.13$, $p < .001$, and interaction between gender and association of domain, $F(1,21098) = 314.39$, $p < .001$. Holm-adjusted pairwise contrasts showed that when painful scenario were typical of men (masculine), participants believed that Black women would feel significantly more pain than Black men, $t(21098) = -6.96$, $\underline{p} < .001$. However, when the type of painful scenario was typical of women (feminine), participants believed that Black women would feel significantly less pain than Black men, $t(21098) = 18.11$, $p < .001$.

Figure 9

Pain Perception Ratings of Black Men and Women



Note. Error bars represent the standard error of the mean.

Prototypicality and Hardship

Race Prototypicality

Supporting H1-B, that darker-skinned people would be perceived as more prototypical than lighter-skinned people, results revealed a main effect of skin tone, $F(1,2040) = 1642.66$, $p < .001$, main effect of gender, $F(1,2040) = 94.044$, $p < .001$ and an interaction between skin tone and gender, $F(1,2040) = 181.59$, $p < .001$. Holm-adjusted pairwise contrasts investigating the interaction showed that participants found darker-skinned Black men and women vs. lighter-skinned as more prototypical of their race, DM-LM: $t(2040) = 40.53$, $p < .001$; DW-LW: $t(2040) = 21.47$, $p < .001$. Skin tone again had a larger impact on how prototypical Black men were viewed of their race, $t(2040) = 30.94$, $p < .001$. This replicated the prototypicality results of Study 2.

Gender Prototypicality

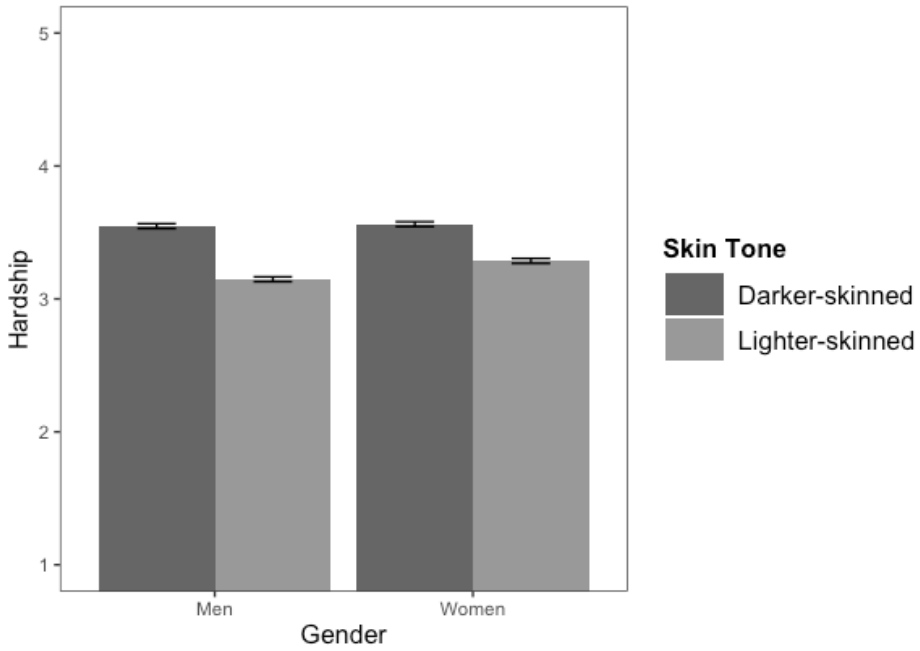
Supporting H1-C, H2-B and replicating the prototypicality results of Study 2, results revealed main effects of skin tone and gender, $F(1,2040) = 59.65$, $p < .001$; $F(1,2040) = 579.092$, $p < .001$, and an interaction, $F(1,2040) = 208.618$, $p < .001$. Holm-adjusted contrasts confirmed that participants viewed Black women as less representative of their gender overall and particularly so if they had dark skin: DW-LW: $t(2040) = -12.70$, $p < .001$. Conversely, Black men were seen as more representative of their gender, and most representative if they had dark skin, DM-LM: $t(2040) = 7.72$, $p < .001$. Skin tone had a larger impact on how prototypical Black women were seen of their gender compared to Black men $t(2040) = -16.341$, $p < .001$.

Assumed Hardship

Results showed that as expected and supporting H1-D, darker-skinned people were thought to have endured more hardship than lighter-skinned Black people, $F(1,2040) = 444.63$, $p < .001$. There was also an interaction, $F(1,2040) = 21.16$, $p < .001$. Holm-adjusted pairwise contrasts investigating the interaction showed that the difference in assumed hardship for darker- vs. lighter-skinned Black people was larger for Black men, DM-LM: $t(2040) = 21.09$, $p < .001$; DW-LW: $t(4020) = 14.58$, $p < .001$. In other words, skin tone had a larger impact on assumed hardship for Black men compared to Black women, $t(2040) = 21.86$, $p < .001$. See Figure 10.

Figure 10

Assumed Hardship by Gender and Skin Tone



Note. Error bars represent the standard error of the mean.

Prototypicality and Hardship as Mediators of Colorism in Pain Perception

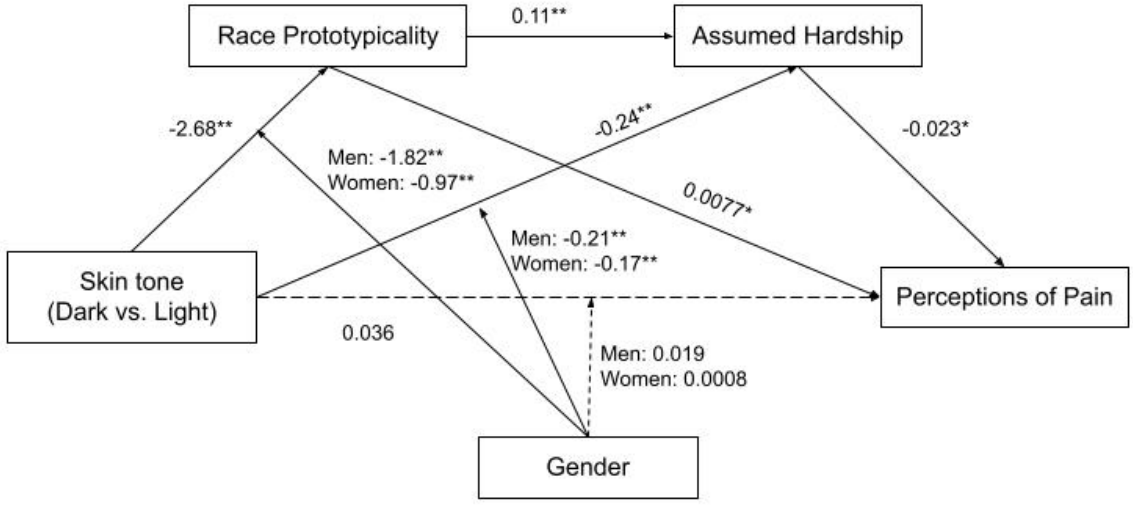
To test H1-E, I first tested the interaction of hardship and racial prototypicality on pain perception. Results showed an interaction between assumed hardship and racial prototypicality, $F(1,24753) = 11.04$, $p < .001$. Post-hoc holm-adjusted pairwise contrasts showed that as levels of assumed hardship increased, participants believed that Black men and women at the mean and +1 SD above the mean for racial prototypicality felt less pain, $t(29085) = 4.45$, $p < .001$.

I then conducted a serial moderated-mediation model using the PROCESS macro (Hayes, 2013), model 85, using skin tone as a predictor variable, prototypicality of race and assumed hardship as mediators and gender as a moderator, controlling for domain. The bootstrap analysis yielded a 95% confidence interval that did not include 0, *95% CI* = [-0.0036, -0.0006], suggesting that gender moderated the mediation effects of prototypicality of race and assumed

hardship on pain perception, *index of moderated mediation* = -0.0021, *SE* = 0.0007. The model was significant for both levels of the moderator (men and women), and conditional indirect effect for the full model was strongest for Black men, *effect* = 0.0044, *SE* = 0.0016, 95% *CI* [0.0013, 0.0076], and weakest for Black women, *effect* = 0.0023, *SE* = 0.0008, 95% *CI* [0.0007, 0.0040], and the pairwise contrast between these effects was significant, -0.0021, *SE* = .001, 95% *CI* [-0.004, -0.001]. See Figure 11.

Figure 11

Moderated-Mediation Model of Skin Tone's Relationship to Perceptions of Pain



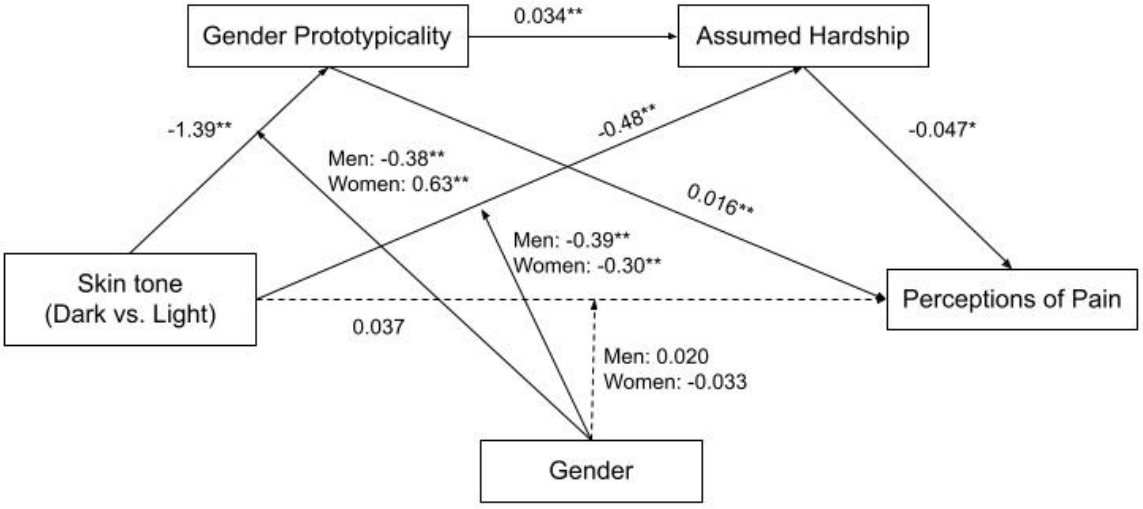
Note. $p < .05$*, $p < .001$**

Gender-Salient Pain Scenario Mediation. To analyze the mediation and interaction in gender-salient scenarios proposed in H2-C, I again conducted a serial moderated-mediation model using the PROCESS macro (Hayes, 2013) model 85, using skin tone as a predictor variable, prototypicality of gender and assumed hardship as mediators and gender as a moderator. The bootstrap analysis yielded a 95% confidence interval that did not include 0, 95% *CI* = [-0.0025, -0.0008], suggesting that gender moderated the mediation effects of prototypicality of gender and assumed hardship on pain perception, *index of moderated*

mediation = -0.0016, *SE* = 0.0004. The model was significant for both levels of the moderator (men and women), and conditional indirect effect for the full model was strongest for Black women, *effect* = -0.0010, *SE* = 0.0003, 95% CI [-0.0015, -0.0005], compared to Black men, *effect* = 0.0006, *SE* = 0.0003, 95% *CI* [0.0003, 0.0010]: the pairwise contrast between these effects was significant, -0.0016, *SE* = .004, 95% *CI* [-0.0025, -0.0008]. See Figure 12. This result was consistent with the prediction of H2-C, that prototypicality of gender will mediate the relationship between skin tone and pain ratings to a greater extent for Black women when gender was salient.

Figure 12

Moderated-Mediation Model of Skin Tone's Relationship to Perceptions of Gender-Salient Pain



Note. $p < .05$*, $p < .001$**

Neutral Pain Scenario Mediation. To test whether gender prototypicality and assumed hardship uniquely mediate the relationship between skin tone and pain perception for gender-salient scenarios, I conducted the same moderated-mediation model using the PROCESS macro (Hayes, 2013) model 85 for neutral pain scenarios. The bootstrap analysis yielded a 95% confidence interval that did include 0, *95% CI* = [-0.0012, 0.0028], suggesting that gender did not

moderate the mediation effects of prototypicality of gender and assumed hardship on pain perception, *index of moderated mediation* = -0.0008, *SE* = 0.0010. The moderated-mediation was only significant for gendered scenarios, and this provided additional evidence that it is the salience of gender that elicits gendered colorism.

Discussion

The results of Study 3 did not show direct condition effects of skin tone on pain perception, which did not support H1-A. However, H1-B, C, D, and E were supported by the data, and racial and gender prototypicality and assumed hardship predicted bias in pain perception. Participants rated darker-skinned Black people as more prototypical of their race, assumed they would have experienced more life hardship, and both racial prototypicality and assumed hardship mediated the relationship between skin tone and pain ratings. This indirect pathway was significant.

Similarly, results did not support H2-A, which predicted that in the gender-salient pain scenarios, skin tone would influence pain ratings to a greater extent for Black women. However, H2-B and C were supported by the data. When evaluating gender prototypicality, participants used skin tone to a greater extent for Black women than Black men, and participants found darker-skinned Black women as the least prototypical of their gender. Importantly, these results replicated the prototypicality results of Study 2. Lastly, in the gender-salient pain scenarios, prototypicality of gender mediated the relationship between skin tone and pain ratings, and this effect was larger for Black women than Black men, meaning that prototypicality of gender held greater weight for participants' evaluations of them. For neutral scenarios, there was no moderation by gender, affirming that this relationship was unique to gender-salient scenarios. Here, we can again see how gendered contexts give rise to gendered colorism for Black women.

General Discussion

Colorism is form of racism that continues to affect the lives of Black Americans, their experiences, and how they are perceived by others. This form of discrimination is both pervasive and dangerous: evidence shows that even children as young as seven years old associate negative stereotypes with darker-skinned peers (T. L. Williams & Davidson, 2009), and darker skin color even within families predicts higher incidences of developing hypertension (Laidley et al., 2019). For Black women and girls, experiences of colorism are related to poorer self-esteem (Wilder & Cain, 2011), and colorist discrimination affects them more prominently in domains such as school punishment (Blake et al., 2017) and media representation (Alter et al., 2016). As Black women face unique struggles related to their intersecting identities, it is both important and necessary to investigate how, when, and why colorism affects Black women in various contexts, experimentally and practically. Across three studies, I explored how and when the perceived non-prototypicality of Black women leads to gendered colorism and connected this bias to pain perception. I proposed that in categorization and evaluation there would be skin tone bias, and that these biases would be more extreme for Black women in domains where gender is salient.

Skin Tone Bias

Studies 1a & b showed a consistent skin tone bias in both racial and gender categorization. Darker-skinned Black men and women were categorized as Black more quickly and accurately than lighter-skinned Black people. Darker skin tones facilitated the fast and accurate categorization of Black men as men but hindered the quick and accurate categorizations of Black women as women. Study 2 moved from skin tone biases in categorization to evaluation, providing evidence that people use skin tone to evaluate prototypicality of race and gender. Participants saw darker-skinned people as more prototypical of Black. Participants also used skin

tone as a feature of gender; participants viewed Black women with darker skin as least prototypical of their gender and Black men with darker skin as most prototypical of their gender. Study 2 also showed that skin tone influenced judgments of similarity. Participants categorized people primarily by race and skin tone.

Lastly, Study 3 connected skin tone biases within evaluations to pain perception, a consequential area of healthcare with well-documented evidence of racism in pain perception and treatment. I did not find direct evidence of skin tone bias in pain perception; skin tone did not influence participants' evaluations of how painful everyday scenarios would be for Black men and women. However, I did find skin tone biases in prototypicality of race and gender, which replicated Study 2 prototypicality results. Darker skin again led to greater perceptions of racial prototypicality and assumed hardship for Black men and women and gender prototypicality for Black men, but lower perceptions of gender prototypicality for Black women. Both assumed hardship and prototypicality also influenced perceptions of pain, mediating the relationship between skin tone and pain perception. Observers did not directly link darker skin to feeling less pain, but the implications of this finding are that skin tone influences multiple processes that pain bias arises from.

Gendered Colorism

Above and beyond skin tone bias, I predicted that colorism would be gendered when gender is made salient, meaning that Black women would incur greater penalties for their skin tone relative to Black men. Because race is a gendered construct, I hypothesized that skin tone biases would be more pronounced for Black women in both racial and gender categorization. Study 1 showed that gendered skin tone biases were only consistently present for gender categorization, suggesting that assumed gender identity and expectations must be salient in a task

for colorism to become gendered as well. I found a consistent gendered skin tone bias, such that dark skin facilitated the categorization of Black men as men but hindered the categorization of Black women as women, and darker-skinned Black women were categorized the slowest and least accurately. In Study 2, participants rated Black women as less prototypical of their gender than White women, particularly if they had dark skin, while dark skin led participants to see Black men as more prototypical of their gender. In this task, gender was salient, and the observed skin tone bias was greater for Black women compared to Black men, consistent with expectations and Study 1 categorization results. After priming gender salience in Study 2, I did not find the expected differences in how participants rated the similarity of Black and White men and women in the gender condition vs. neutral control condition. Instead, participants consistently rated similarity according to race and skin tone, with darker-skinned Black men and women visualized as the least similar to White men and women. Participants saw Black women as less "woman" than White women, particularly if they were darker-skinned. Contrarily, for Black men, darker skin was seen as more "man". Here, the association of "Blackness" with masculinity led to greater skin tone bias against Black women and erasure of their gender identity.

Lastly, in Study 3, I did not find direct evidence of gendered colorism in ratings of pain for everyday gender-salient scenarios vs. neutral scenarios. However, there was again evidence of gendered colorism in ratings of gender prototypicality, such that Black women incurred greater erasure of their gender because of their skin tone, replicating the results of Study 2. Among gender-salient painful scenarios, prototypicality of gender and assumed hardship both predicted perceptions of pain, and the indirect effect was stronger for Black women relative to Black men. Here, I found indirect evidence of gendered colorism, as skin tone affected

perceptions of gender and assumed hardship, which in turn predicted pain ratings. The effects of gendered colorism on pain perception are complex and tied to gender salience for Black women. Thus, the way skin tone bias may affect perception and treatment of Black women's pain in "real life" is not so straightforward, and addressing this bias requires understanding how multiple forms of oppression interact.

Implications

These results provide further evidence that colorism affects Black men and women, and disproportionately so for Black women when gender stereotypes, expectations, or assumed identity are centrally relevant. In Study 1, this occurred when participants were slower and less accurate at categorizing Black women as women, particularly when they had dark skin. In Study 2, this was true when participants evaluated Black women's prototypicality of gender based on skin tone. In Study 3, this occurred again in participants' ratings of Black women's gender prototypicality, which in turn influenced assumptions of life hardship and perceptions of pain in gendered scenarios. These insights extend the existing literature on skin tone bias. It adds nuance to our understanding of how and when colorism becomes gendered. This data suggests that skin tone biases can manifest in harmful ways, in ways that at times are more for Black men, and when gender is salient, for Black women.

This work also challenges some theories about racism. The subordinate male target hypothesis (SMTH) (Sidanius & Pratto, 1999) posits that Black males are the primary targets of racism, experiencing much more than their female counterparts. There is research to support this theory, such as employment discrimination audit studies and self-reports of daily discrimination in the workplace (Sidanius & Pratto, 1990). However, this account does not consider the effect that skin tone has on how Black people experience racist discrimination. Colorism intersects with

racial and ethnic oppression, greatly impacting how Black people experience racism as well as gendered racism. The present work takes an intersectional approach, providing evidence that both Black men and women can "bear the brunt" of racist discrimination. Moreover, it begins to identify when and why they do: namely, when race vs. gender is salient, and because of perceptions about Black men and women's prototypicality. This can lead to Black women experiencing more direct forms of discrimination. Although there were no direct effects of skin tone on pain ratings on Study 3, the indirect effects of skin tone on racial and gender prototypicality and hardship that lead to lower pain ratings for darker-skinned Black women are direct forms of prejudice. Again, this is also evidenced in school punishment outcomes (Hannon et al., 2013; Blake et al., 2017) and the darkening of Black women's skin in negative media stories (Alter et al., 2016).

The present work fits into the framework of the Intersectional Invisibility hypothesis—which analyzes how multiple marginalized identities intersect with another and influence how forms of oppression affect them. The resulting experiences differ from people with a single marginalized identity (Purdie-Vaughns & Eibach, 2008). As previously noted, this hypothesis suggests that because of Black women's perceived non-prototypicality of their race and gender, Black women often experience discrimination in forms of neglect and invisibility. This intersectional invisibility is defined as "the general failure to fully recognize people with intersecting identities as members of their constituent groups" (Purdie-Vaughns & Eibach, 2008). In the present work, the slowness and especially inaccuracy of recognition of Black women's gender are failures to fully recognize them as women. The consistent devaluation of Black women's gender prototypicality, particularly darker-skinned Black women's, is a denial of part of their identity. This erasure of their gender is a form of intersectional invisibility. I argue

then, that analyzing how skin tone intersects with race and gender provides evidence that Black women experience both direct and passive forms of discrimination due to the same characteristics—their perceived non-prototypicality. Gendered colorism is not as simple as just skin tone—it is a manifestation of the intersection of multiple forms of oppression.

This work can also provide insight on how colorism manifests in the real world, particularly when and why Black women may face the brunt of it. As previously noted, Black women and girls are currently known to face more colorist discrimination in domains like school punishment (Blake et al., 2017; Hannon et al., 2013), dating and marriage (Hunter, 2007), and media portrayal (Alter et al., 2016). Notably, these are also domains in which gender stereotypes and expectations are salient and centrally relevant to judgments. Accordingly, this research can also help us predict and study additional areas where this form of discrimination may particularly hurt Black women, such as pregnancy and childbirth outcomes, and other gender-salient domains where racial inequities in healthcare and access exist. As evidenced in Study 3, it is not simply areas traditionally thought of as women's health that can trigger gendered colorism in pain perception, but any scenario in which assumed gender identity, stereotypes, or expectations may still be activated (e.g. burning oneself while *baking*, shoulder pain from carrying a heavy *purse*).

The indirect effects of skin tone bias on pain perception also have notable implications for our existing knowledge on pain perception and treatment. Essentially, the study revealed multiple pathways for pain bias occurring. Gender prototypicality, racial prototypicality, and assumed hardship are processes by which skin tone influences pain bias. It is not skin tone alone that may lead to providers underestimating the pain their Black patients feel, but what skin tone represents. Essential to skin tone bias is perceived "Blackness", and the stereotypes of masculinity and toughness that come with it. Given that the effects of skin tone on pain bias were

nuanced, so too must be the solution. To ameliorate this issue of pain bias against darker-skinned Black women and Black people as a whole, solutions must focus on eliminating the pathways by which skin tone influences pain bias, rather than simple education on skin tone bias alone.

Additional to these insights on gendered colorism in the real world, Study 3 revealed that Black women were believed to experience less pain relative to Black men in gendered domains, specifically gendered pain associated with women. These results are important in context with how women's pain is treated broadly. Decades of research shows that women's complaints of pain are often discounted (D. E. Hoffmann & Tarzian, 2001) and on a larger scale, women's health issues are underfunded (Mirin, 2021). More so, when diseases affect one gender more than another, research on diseases that affect women more often than men are underfunded, while research on diseases that affect men more often than women are overfunded (Mirin, 2021). These results show that this broad trivialization of women's pain in healthcare goes as far as trivializing women's pain in everyday life.

Limitations

The current studies also present their own limitations. Study 1 used reaction time data. And although these are widely used to measure cognitive processing and, specifically, ease of categorization, there are limitations to the use of reaction times. Slow reaction time can reflect lapses in attention (De Boeck & Jeon, 2019), or even increased attention. For these reasons, I analyzed both accuracy and reaction time. Black women were categorized more slowly as women, but darker-skinned Black women were categorized both more slowly and inaccurately, providing more confidence these data reflect a skin tone bias in gender categorization, specifically a denial of darker-skinned Black women's gender.

In Study 2, I used a Multidimensional Scaling paradigm. Although Multidimensional Scaling allows for visualizations of data that provide insight into how people categorize or understand items, concepts, or people, there are limitations to using pairwise similarity ratings. These types of similarity tasks necessitate a smaller number of items as stimuli because each additional item greatly increases the duration of the task for participants. In Study 2, participants viewed two faces of each gender and skin tone combination, resulting in 66 item-by-item comparisons. A larger amounts of items per group would increase the generalizability of the results.

Moreover, in Studies 2 and 3, I relied on self-report measures, which are susceptible to desirability concerns. These concerns may be especially salient in within-participants design, such as the one I used. Indeed, previous studies on racial bias in pain perception have used mostly between-subject designs, asking participants to evaluate either a Black or White target person, but not both. Doing this can reduce desirability concerns (Druckman et al., 2018; Trawalter et al., 2012; cf., Hoffman et al., 2016). To maximize power and because my previous studies showing colorism have also been within-subjects design, I used a within-participant design. I did not ask participants about their self-presentational concerns (e.g., motivation to appear non-prejudiced). Replicating Study 3 using a between-participant design and/or with a measures of self-presentation concerns, such as the motivation to appear non-prejudiced, could provide insight. It could help us determine whether the null effect of skin tone (and skin tone and gender) on pain ratings might be due to self-presentational concerns.

In addition, the scenarios in Study 3 were hypothetical. Participants reported how they would anticipate feeling and how another person would feel if a painful scenario occurred instead of reporting pain from actual experiences. This presents a limitation, as people often

underestimate the intensity of future physical pain (Read & Loewenstein, 1999), and are not always accurate at predicting their future behavior. Depending on how proximal or distal an event or scenario is, people also use different contextual factors to predict their future behavior and are more accurate the more proximal said event is to occur (Persky et al., 2007). Study 3 results reflect biases in predicted behavior, but do not show responses to real-life events.

Limitations aside, these studies provided unique and important evidence of how and when colorism becomes gendered.

Future Directions

The present studies and their limitations open multiple avenues for future research. Skin tone was more consequential for men when race was salient in the evaluation task for Studies 2 and 3, but not Study 1. Future studies should investigate this effect to see if there is a consistent elevated skin tone bias for Black men vs. Black women when race is of primary salience, or if there are secondary factors (i.e., facial features) that would lead to a gendered skin tone bias for men vs. women. For example, facial features and skin tone only moderately covary, and research shows that Afrocentric facial features (i.e., broader nose and or fuller lips), both independently and jointly with skin tone, elicit biased impression judgments from others (Hagiwara et al., 2012). Future research could investigate how colorism may be magnified or reduced depending on how Afrocentric one's facial features may be.

Future studies considering MDS or similar visualization maps could employ methods such as the spatial arrangement method (SpAM) (Goldstone, 1994) in order to have a shorter task with a larger number of faces as stimuli. In this method, participants see all items at once and arrange them in order of similarity to each other. While the number of items presented can still present a demanding task for participants, the SpAM method can be completed in a shorter

amount of time with a larger number of stimuli. Similarly, researchers employing future MDS studies could create dissimilarity scores based on a gender-salient question, to ensure that gender is salient.

Additionally, future research can also examine more factors that predict skin tone bias and gendered skin tone bias for Black women. Current research provides evidence of multiple predictors of racialized pain bias: assumed hardship (Druckman et al., 2018), biological beliefs about differences between Black and White bodies, (K. M. Hoffman et al., 2016) and racial prototypicality (Drain et al., 2020) all predict racism in pain perception. Future studies investigating unique predictors of gendered racism and gendered colorism in pain perception could provide necessary information on how and when women would be most at risk at experiencing these biases in healthcare settings. Lastly, as the effect of skin tone on pain bias was correlational, it is important and necessary for future studies to investigate whether there is direct and casual evidence of colorism and gendered colorism in pain perception. As noted above, future studies then on skin tone pain perception biases should measure motivations to appear non-prejudiced to assess the potential issue of desirability concerns.

Future studies then should also include audit studies of colorism in pain perception to assess whether and how colorism presents a problem currently in this domain. Specifically, these studies could investigate health domains such as chronic pain, which women are more likely to suffer from (Bartley, & Fillingim, 2013). This would address several limitations, including the limitations of self-reported future behavior. Considering that laypeople are not always able to accurately predict their future behavior or their responses to future physical pain, future studies should include medical personnel evaluating medical scenarios similar to those they evaluate in

real-life on a daily basis, as their predictions of how they would treat their patients would likely reflect real-life biases.

The present studies had a multiracial participant pool. Even so, most participants were White, so it was not possible to analyze the data for each group of color represented in the data. An important future direction for studies on colorism and gendered colorism are to investigate the psychology behind these biases for Black Americans especially, as well as other racial and ethnic groups with legacies of colorism. Although in-group and out-group colorism occur in the real world and are documented experimentally, this does not necessarily mean that the consequences of, motivations, and psychological processes behind within-group colorism are the same in an out-group colorism context. Future studies should investigate whether Black participants show the same gendered skin tone biases in categorization, and if so, the psychological processes behind it.

Concluding Remarks

Altogether, this research contributes to the growing body of literature focusing on and acknowledging the ways that Black women uniquely experience racialized and gendered oppression. Like race and gender, skin tone has structural qualities that make colorism a necessary part of researching race and racism. I conclude this paper with a quote by Alice Walker, who coined the term colorism in her 1981 essay, *If the Present Looks Like the Past, What Does the Future Look Like?*:

> The problem of the twenty-first century will still be the problem of the color line, not only "the relation of the darker to the lighter races of men [sic] in Asia and Africa, in America and the islands of the sea," but the relations between the darker and lighter people of the

> same races, and of the women who represent both dark and light within each race…it is the whole family, rather than the dark or the light, that must be affirmed (p. 311).

40 years later, Alice Walker's words remain prophetic. Colorism remains a persistent problem in the 21st century, and its consequences for the lives and experiences of Black women are important and worthy of attention and care.

References

Allure. (2020, January 28). *100 Years of Iconic Hairstyles* [Video]. YouTube. https://www.youtube.com/watch?v=PMrSzK2qvuY&t=2s

Alter, A. L., Stern, C., Granot, Y., & Balcetis, E. (2016). The "Bad Is Black" Effect. *Personality and Social Psychology Bulletin*, 42(12), 1653-1665.

Anderson, K. O. et al. (2009). Racial and Ethnic Disparities in Pain: Causes and Consequences of Unequal Care. *The Journal of Pain*, 10(12), 1187-1204.

Barsky, A. J., Peekna, H. M., & Borus, J. F. (2001). Somatic symptom reporting in women and men, *Journal of General Internal Medicine*, 16(4), 266–275.

Bart-Plange, D., & Trawalter, Sophie (advisor) (2019). *Colorism in Context: Exploring Differential Effects of Gender on Skin Tone Bias*. Charlottesville, VA: University of Virginia, Psychology - Graduate School of Arts and Sciences, MA (Master of Arts), 2019. Retrieved from https://doi.org/10.18130/v3-matp-kr56.

Bartley, E. J., & Fillingim, R. B. (2013). Sex differences in pain: a brief review of clinical and experimental findings. *British Journal of Anaesthesia*, 111(1), 52–58. https://doi.org/10.1093/bja/aet127

Beach, M. C., Saha, S., Park, J., Taylor, J., Drew, P., Plank, E., Cooper, L. A. & Chee, B. (2021). Testimonial injustice: linguistic bias in the medical records of black patients and women. *Journal of General Internal Medicine*, *36*(6), 1708-1714.

Crowston K. (2012) Amazon Mechanical Turk: A Research Tool for Organizations and Information Systems Scholars. In: Bhattacherjee A., Fitzgerald B. (eds) Shaping the Future of ICT Research. Methods and Approaches. IFIP Advances in Information and

Communication Technology, vol 389. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35142-6_14.

De Boeck P and Jeon M (2019) An Overview of Models for Response Times and Processes in Cognitive Tests. *Frontiers in Psychology*. 10(102). doi: 10.3389/fpsyg.2019.00102

DeBruine, L. (2018). debruine/webmorph: Beta release 2 (Version v0.0.0.9001). Zenodo. http://doi.org/10.5281/zenodo.1162670

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language*, 59 (2008) 390–412. doi:10.1016/j.jml.2007.12.005

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, 15, 674–679.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology*, 83, 5–25.

Blake, J. J., Keith, V. M., Luo, W., Le, H., & Salter, P. (2017). The Role of Colorism in Explaining African American Females' Suspension Risk. *School Psychology Quarterly,* 32(1), 118-130. http://dx.doi.org/10.1037/spq0000173.

Cleeland, C. S., Gonin, R., Baez, L., et al. (1997). Pain and Treatment of Pain in Minority Patients with Cancer: The Eastern Cooperative Oncology Group Minority Outpatient Pain Study. *Annals of Internal Medicine*, 127:813–816.

Collins, P. H. (1990). *Black Feminist Theory: Knowledge, Consciousness, and the Politics of Empowerment*. Hyman.

Cooley, E., Winslow, H., Vojt, A., Shein, J., & Ho, J. (2018). Bias at the intersection of identity: Conflicting social stereotypes of gender and race augment the perceived femininity and interpersonal warmth of smiling Black women. *Journal of Experimental Social Psychology*, 74, 43-49.

Drain A, Goharzad A, Qu-Lee J, Lin J, Mende-Siedlecki P. (2020) Racial prototypicality cues exacerbate racial bias in pain perception from the bottom up and the top down. PsyArXiv;. DOI: 10.31234/osf.io/6gu8j.

Druckman, J. N., Trawalter, Montes, S. I., Fredendall, A., Kanter, N., & Rubenstein, A. P. (2018) Racial bias in sport medical staff's perceptions of others' pain, *The Journal of Social Psychology,* 158:6, 721-729, DOI: 10.1080/00224545.2017.1409188.

Eberhardt, J.L., Davies, P.G., Purdie-Vaughns, V.J. and Johnson, S.L. (2006). 'Looking Deathworthy: Perceived Stereotypicality of black defendants predicts capital- sentencing outcomes', *Psychological Science*, 17(5), pp. 383–386.

Faul F., Erdfelder E., Buchner A., & Lang A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*. 41(4):1149-1160. DOI: 10.3758/brm.41.4.1149. PMID: 19897823.

Feliciano, C. (2016). Shades of race: How phenotype and observer characteristics shape racial classification. *American Behavioral Scientist*, 60, 390-419.

Fisk, S. T., & Neuberg, S. (1990). A Continuum of Impression Formation, from Category-Based to Individuating Processes: Influences of Information and Motivation on Attention and Interpretation. *Advances in Experimental Social Psychology*, *23*(C), 1-74. https://doi.org/10.1016/S0065-2601(08)60317-2

Funk, C., (2022, April 7). *Black Americans' Views of and Engagement With Science*. Pew Research Center. https://www.pewresearch.org/science/2022/04/07/black-americans-views-of-and-engagement-with-science/.

John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Greenwald, A. G., Nosek, B. A., Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. Attitudes and Social Cognition. 85(2), 197-216, DOI: 10.1037/0022-3514.85.2.197.

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, 392–403.

Goldsmith, A. H. & Hamilton, D. & Darity, W. (2007). From Dark to Light: Skin Color and Wages Among African-Americans, *Journal of Human Resources*, University of Wisconsin Press, vol. 42(4).

Goldstone R. An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers.* 1994;26:381–386.

Goyal, M. K., Kupperman, N., Cleary, S. D., Teach, S. J., & Chamberlain, J. M. (2015). Racial disparities in pain management of children with appendicitis in emergency departments. *JAMA Pediatrics*, 169, 996–1002. doi:10.1001/jamapediatrics.2015.1915

Johnson, L. K., Freeman, B. J., & Paulker, K. (2012). Race is Gendered: How Covarying Phenotypes and Stereotypes Bias. *Journal of Personality and Social Psychology*, 102(1), 116-131.

Johnson, P. A., Golman, L., & Orav, E. J. (1996). Gender Differences in the Management of Acute Chest Pain: Support for the "Yentl" Syndrome. *Journal of General Internal Medicine*, 11, 209-17.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Hagiwara, Nao & Kashy, Deborah & Cesario, Joseph. (2012). The Independent Effects of Skin Tone and Facial Features on Whites' Affective Reactions to Blacks. *Journal of Experimental Social Psychology*. 48. 892. 10.1016/j.jesp.2012.02.001.

Hall, J. C. (2017). No Longer Invisible: Understanding the Psychosocial Impact of Skin Color Stratification in the Lives of African American Women. *Health & Social Work,* 42(2), 71-78. doi: 10.1093/hsw/hlx001.

Hannon, L., Defina, R., & Bruch, S. (2013). The Relationship Between Skin Tone and School Suspension for African Americans. *Race and Social Problems*, 5(4), 281-295.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hoffman, D. E. & Tarzian, A. J. (2001). The Girl Who Cried Pain: A Bias Against Women in the Treatment of Pain. J*ournal of Law, Medicine & Ethics*, 29 (2001): 13–27.

Hoffman, K. M. & Trawalter, S. (2016). Assumptions about life hardship and pain perception. *Group Processes & Intergroup Relations,* 19(4), 493-508.

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

Holden, K. B., Bradford, L. D., Hall, S. P., & Belton, A. S. (2013). Prevalence and correlates of depressive symptoms and resiliency among African American women in a community-based primary health care center. *Journal of Health Care for the Poor and Underserved*, 24(4 Suppl), 79–93. https://doi.org/10.1353/hpu.2014.0012

hooks, b. (1987). *Ain't I a Woman.* Pluto Press.

Hout, M. C., Goldinger, S. D., & Brady, K. J. (2014). MM-MDS: a multidimensional scaling database with similarity ratings for 240 object categories from the Massive Memory picture database. *PloS one*, *9*(11), e112644. https://doi.org/10.1371/journal.pone.0112644

Hughes, M., & Hertel, B. R. (1990). The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans. *Social Forces*, 68(4), 1105. doi:10.2307/2579136.

Hunter, M. L. (2002). "If You're Light You're Alright": Light Skin Color as Social Capital for Women of Color. *Gender & Society*, 16(2), 175–193. https://doi.org/10.1177/08912430222104895

Hunter, M. (2005). *Race, gender, and the politics of skin tone.* New York, NY: Routledge.

Hunter, M. L. (2007). The Persistent Problem of Colorism: Skin Tone, Status, and Inequality. *Sociology Compass*, 1(1), 237-254.

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*(1), 54–69. https://doi.org/10.1037/a0028347

Laidley, T., Domingue, B., Piyapat Sinsub, P., Mullan Harris, K. & Dalton Conley, D. (2019). New Evidence of Skin Color Bias and Health Outcomes Using Sibling Difference Models: *A Research Note. Demography*, 56(2), 753-762.

Logé, C., Ross, E., Dadey, D. Y. A., Jain, S., Saporta, A., Ng. A. Y., & Pajpurkar, P. (2021). Q-Pain: A Question Answering Dataset to Measure Social Bias in Pain Management. *NeurIPS*. https://doi.org/10.13026/2tdv-hj07

Ma, Correll, & Wittenbrink (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. *Behavior Research Methods*, 47, 1122-1135.

Maddox, K. B., & Chase, S. G. (2004). Manipulating subcategory salience: exploring the link between skin tone and social perception of Blacks. *European Journal of Social Psychology*, 34(5), 533-546.

Maddox, K. B., & Gray, S. A. (2002). Cognitive Representations of Black Americans: Reexploring the Role of Skin Tone. *Personality and Social Psychology Bulletin, 28*(2), 250-259. doi:10.1177/0146167202282010

Massey, D. S., Martin, J. A. (2003). *The NIS Skin Color Scale*. Princeton, NJ: Princeton University Press.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4):630–633, 2004.

Mirin A. A. (2021). Gender Disparity in the Funding of Diseases by the U.S. National Institutes of Health. *Journal of Women's Health,* (2002), 30(7), 956–963. https://doi.org/10.1089/jwh.2020.8682

Monk, E. (2015). The Cost of Color: Skin Color, Discrimination, and Health among African-Americans. *American Journal of Sociology*, 121(2), 396-444.

Oksanen, J., F., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., , Solymos, P., Henry M., Stevens, H., Szoecs, E. & Wagner, H. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan.

Persky, S., Kaphingst, K. A., Condit, C. M. & McBride, C. M. (2007). Assessing hypothetical scenario methodology in genetic susceptibility testing analog studies: a quantitative review. *Genetics in Medicine* 9**,** 727–738. https://doi.org/10.1097/GIM.0b013e318159a344

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional Invisibility: The Distinctive Advantages and Disadvantages of Multiple Subordinate-Group Identities. *Sex Role*s, 59(5-6), 377-391.

Reece, R. L. (2018). Genesis of U.S. Colorism and Skin Tone Stratification: Slavery, Freedom, and Mulatto-Black Occupational Inequality in the Late 19th Century. *The Review of Black Political Economy.* 45(1) 3 –21. DOI: 10.1177/0034644618770761.

Read, D., & Loewenstein, G. (1999). Enduring pain for money: Decisions based on the perception and memory of pain. *Journal of Behavioral Decision Making*, 12(1), 1–17. https://doi.org/10.1002/(SICI)1099-0771(199903)12:1<1::AID-BDM310>3.0.CO;2-V

Reuter, E. B. (1918). *The mulatto in the United States: Including a study of the role of mixed-blood races throughout the world*. Badger.

Russell, K., Wilson, M., & Hall, R. (1992). *The color complex: The politics of skin color among African Americans.* New York, NY: Harcourt Brace Jovanovich.

Ryabov, I. (2013). Colorism and school-to-work and school-to-college transitions of African American adolescents. *Race and Social Problems*, 5(1), 15-27. https://doi.org/10.1007/s12552-012-9081-7

Sabin J. A., & Greenwald A. G. (2012). The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *American Journal of Public Health*, 102(5):988-95.

Samulowitz, A., Gremyr, I., Eriksson, E., & Hensing, G., (2018). "Brave Men" and "Emotional Women": A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain. *Pain Research and Management*, 2018(6358624). https://doi.org/10.1155/2018/6358624

Science Insider. (2019, September 13). *What's Inside A Caterpillar 'Cocoon?'* [Video]. YouTube. https://www.youtube.com/watch?v=uK_iZZ4Bx2o&t=2s

Sesko, A. K., & Biernat, M. (2010). Prototypes of race and gender: The invisibility of Black women. *Journal of Experimental Social Psychology*, 46(2), 356-360. http://dx.doi.org/10.1016/j.jesp.2009.10.016

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Stewart, Q. T., Cobb, R. J., & Keith, V. M. (2018). The color of death: race, observed skin tone, and all-cause mortality in the United States. *Ethnicity & Health*, May 8:1-23.

Strohminger, N., Gray, K., Chituc, V. et al. (2016). *Behavioral Research* 48: 1197. https://doi.org/10.3758/s13428-015-0641-9

Thomas, E. L., Dovidio, J. F., & West, T. V. (2014). Lost in the categorical shuffle: Evidence for the social non-prototypicality of Black women. *Cultural Diversity and Ethnic Minority Psychology,* 20(3), 370–376.

Todd, K. H., Deaton, C., D'Adamo, A. P., & Goe, L. (2000). Ethnicity and analgesic practice. *Annals of Emergency Medicine*, 35, 11–16. doi:10.1016/S0196-0644(00)70099-0.

Townsend, T. G., Neilands, T. B., Thomas, A. J., & Jackson, T. R. (2010). I'm no Jezebel; I am young, gifted, and Black: Identity, sexuality, and Black girls. *Psychology of Women Quarterly*, 34(3), 273-285.https://doi.org/10.1111/j.1471-6402.2010.01574.x.

Uzogara, E. E. & Jackson, J. S. (2016). Perceived Skin Tone Discrimination Across Contexts: African American Women's Reports. *Race and Social Problems*, 8, 147-159.

Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S. Fourth Edition*. Springer, New York. ISBN 0-387-95457-0

Walker, A. (1982). *If the present looks like the past, what does the future look like? In search of our mothers' gardens: Womanist prose*. (pp. 290-291). Harcourt.

Westfall, J. (2016). PANGEA: Power ANalysis for GEneral Anova designs.

Whelan, R. (2010). Effective analysis of reaction time data. *The Psychological Record*, 58(3), 9.

Wilder, J., & Cain, C. (2011). Teaching and learning color consciousness in Black families: Exploring family processes and women's experiences with colorism. *Journal of Family Issues, 32*(5), 577–604. https://doi.org/10.1177/0192513X10390858

Williams, T. L., & Davidson, D. (2009). Interracial and intra-racial stereotypes and constructive memory in 7-and 9-year-old African-American children. *Journal of Applied Developmental Psychology*, 30(3), 366-377.