Black racial phenotypicality: Implications for the #BlackLivesMatter Movement☆

Maire L. O'Hagan [*], Samantha R. Pejic, Jason C. Deska

Psychology Department, Toronto Metropolitan University, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada

ABSTRACT

Black individuals with phenotypically African features tend to experience heightened discrimination and mistreatment. The current research examined how racial phenotypicality and prototypicality effect hate crime reporting metrics and beliefs about who evaluators are represented #BlackLivesMatter. Across five studies ($N = 876$), results indicate that, compared to low racially phenotypic Black targets, high phenotypic targets were seen as more represented by #BlackLivesMatter (Study 1). When depicted as being the victim of a hate crime, high phenotypic targets were deemed more credible and that it was more appropriate for them to report their victimization on the #BlackLivesMatter website compared to their low phenotypic counterparts by White (Study 2a and 2c) and Black participants (Study 2b and 2c). Black (Study 2b and 2c) and White (Study 3) participants showed differences in perceptions of harm following hate crime victimization. Study 3 extended these findings to a separate manipulation of prototypicality and used a more ecologically valid context. These findings provide support for the problematic exclusivity of narrow prototypes by demonstrating their effect on beliefs about who social justice movements represent, and how they influence beliefs about victim reporting metrics.

1. Introduction

The population of Black Americans has been steadily increasing, with an estimated 47.2 million, or 14.2 %, of the United States population self-reported their race as being Black as of 2022 (Pew Research Centre, 2023). The racial and ethnic diversity among Black Americans is vast, with individuals representing diverse regions and cultures. Specifically, the genetic makeup, as well phenotypicality (i.e., physical appearance cues that signify racial prototypicality, Blair et al., 2002, 2004; Maddox, 2004; where racial prototypicality is defined as how representative an individual is of their racial category) of Black Americans varies substantially. Disturbingly, while the Black American population continues to increase, incidents of anti-Black race-based hate crimes are prevalent in the United States. A hate crime is defined as a criminal act that is motivated by an individual's identity or perceived identity such as their race or gender (U.S. Department of Justice, 2023). Race-based hate crimes are the most common form of hate crime in the United States, and in 2022 alone, there were 3421 anti-Black hate crimes reported (U.S. Department of Justice, 2023b). Black Americans report more race-based hate crime victimization than any other minoritized group (U.S.

Department of Justice, 2023b). These statistics likely underestimate the true prevalence of these events due to reporting barriers. Considering the growing Black population in the United States, and the prominence of anti-Black hate crimes, it is imperative to understand how we can protect and find justice for Black victims of race-based hate crimes. Further, considering the diversity among the Black population in the United States, it is integral to assess how hate crime reporting metrics are perceived according to Black racial prototypicality. It is possible that when lighter skin Black Americans experience racism, they might be perceived as less credible when reporting their experience due to expectations of a more stereotypical Black appearance from a victim of anti-Black racism. Thus, one of the primary aims of the current research was to address this gap in the prototypicality literature.

One social justice movement in particular, the #BlackLivesMatter movement (Black Lives Matter, 2024), holds promise for increasing racial equity, representation, and inclusion for Black Americans. The #BlackLivesMatter movement is meant to serve as a crucial support system for the Black American community. Yet, while the #Black-LivesMatter movement celebrates diversity and promotes inclusion among Black Americans, public perceptions regarding who should be

represented by the #BlackLivesMatter movement is not yet understood. Indeed, previous research demonstrates the problematic transference of narrow prototypes to who is included in broad racial movements (Pejic & Deska, 2024). Given the heterogeneity among Black Americans, coupled with the demonstrated differences in skin tone sensitivity between Black and White perceivers (Maddox, 2004), understanding perceptions of inclusion in the #BlackLivesMatter movement from both intra- and intergroup perspectives is imperative to ensure equal access to support services among all Black Americans.

The current work examined how perceptions of hate crime reporting metrics are perceived for Black Americans varying in racial phenotypicality and prototypicality, with a specific focus on perceptions of credibility, psychological harm endured, and how appropriateness it is for individuals to report an anti-Black hate crime to the #BlackLivesMatter website. Further, the current research examined how well people perceive Black Americans varying in phenotypicality to be represented by the #BlackLivesMatter movement.

1.1. Historical roots of racism toward Black Americans

The influence of skin tone differences on social and economic development between Black Americans is long-standing, and traces back to the slave trade (see Keith & Herring, 1991; Monk Jr, 2015; Russell et al., 1992). Problematically, the influence of skin tone differences among Black Americans persists today, cementing the relationship between the skin tone and access to economic and social opportunities. For instance, research shows that compared to lighter skinned Black Americans, darker skinned Black Americans experience worse outcomes in educational attainment, occupation selection, as well as average income (Keith & Herring, 1991). Current research holds that skin tone continues to impact the experiences of Black Americans (Monk Jr, 2021). As Maddox (2004) details, skin tone is one of the most readily available features that we use to make racial distinctions, and the preference for Eurocentric features such as lighter skin tone in individuals means that darker skin Black Americans have been subjected to various forms of prejudice and discrimination throughout history. As one's phenotypicality deviates from the norm (i.e., White), targets tend to be evaluated more harshly and such evaluations are apparent in both ingroup and outgroup members' assessments of non-prototypic others. Accordingly, when lighter skin Black Americans experience racism, they might be perceived as less credible when reporting their experience because people expect a more stereotypical Black appearance (i.e., darker skin tone) from a victim of anti-Black racism.

1.2. Prototype of Black Americans

Rosch's (1978) seminal work in cognitive linguistics laid the foundation for social psychological applications of prototype research. We use heuristics, especially those based on visual cues, when classifying stimuli to reduce cognitive load. One common strategy used to quickly and easily classify stimuli is through the use of prototypes, or in other words by considering how well the new stimuli represents the most common example of that category (i.e., the prototype of that category). Integral to Rosch's original proposal of prototypes is the fluidity of what is considered "average" or "prototypical." However, contemporary understandings of social prototypes demonstrate a resistance to sociocultural changes. Not only do we tend to categorize individuals according to how well they align with our ideas of prototypical representations, but we also tend to prefer prototypical members over non-prototypical members (Vogel et al., 2020). Critically, our use of prototypes to classify others is not benign.

For example, Goh et al. (2021) examined the downstream consequences of gender prototypicality by examining how adherence to prototypes affect outcomes associated with reporting sexual harassment. Through 11 studies, participants' mental representations of sexual harassment victims were highly similar to that of prototypical women,

participants were less inclined to deem sexual harassment incidents as such for non-prototypical women compared to prototypical women, sexual harassment claims were deemed less credible, and perceivers believed victims endured less psychological harm when the women were non-prototypical compared to prototypical. Although these findings focus on being non-prototypical of one's gender, they suggest that non-prototypical victims of crimes related to atypical identity may be at a disadvantage in terms of their perceived believability and harm endured as a victim.

Most germane to the current project, skin tone is a dominant appearance cue influencing prototypes (e.g., Maddox, 2004). Skin tone biases affect how we both perceive and interact with others, and skin tone bias has been shown to be activated in both Black and White individuals (Maddox & Gray, 2002). Indeed, most prototypicality research on Black Americans focuses on skin tone, with lighter skinned individuals in North America being perceived as less prototypical of the Black racial category due to associations with European ancestry. Research suggests that darker skin is linked to higher perceived African[1] prototypicality, as seen in research assessing the influence of facial metrics and skin tone on prototypicality evaluations (Strom et al., 2012). Notably, this effect varies according to the evaluator's race. White perceivers are less sensitive to differences in skin tone than Black perceivers, suggesting that skin tone differences may not impact White observers' racial categorizations to the same extent as racial ingroup members (Maddox, 2004). These findings underscore the importance of examining evaluations from both intra- and inter-group members.

One way that skin tone bias works is through stereotype activation (Adams et al., 2016). Racial stereotypes are more strongly activated for those we perceive to be better representative members of a racial category (i.e., most prototypical; Blair et al., 2002). In a study assessing the activation of racial stereotypes, participants reported more negative traits for dark skinned Black targets compared to their lighter skinned Black counterparts (Maddox & Gray, 2002). They further reported significantly more positive traits for lighter skinned Black targets, demonstrating a greater association of positive traits with lighter skinned Black targets compared to dark skinned targets. Other work finds that more prototypically Black targets appear angrier, more dominant, more masculine, and tougher than less prototypically Black targets, an effect that influences beliefs about how sensitive people are to pain, and consequently, how much support they should receive (Deska et al., 2020). Black racial prototypicality affects implicit and explicit prejudice (Livingston & Brewer, 2002), and has important implications for criminal sentencing. Those with more pronounced prototypical features are given more severe sentences, including the death penalty, compared to their less prototypical counterparts (Blair et al., 2004; Eberhardt et al., 2006). Taken together, those with more prototypical Black features tend to be ascribed Black stereotypes more strongly, and consequently experience heightened discrimination. Accordingly, their experiences of racism may be more readily recognized and validated by others due to their prototypicality of the Black racial category. Considering the diversity in phenotypicality among Black Americans, it is imperative to explore how the experiences of prototypically diverse victims of anti-Black hate crimes are understood and interpreted to ensure equal access to support services intended for all Black Americans. Further, it is important to examine how Black racial diversity maps onto who the public perceives are best represented by racial social justice movements that aim to put an end to racial injustice and discrimination, such as #BlackLivesMatter.

1.3. The #BlackLivesMatter movement

#BlackLivesMatter is a social justice movement through which the

---

[1] Strom et al. (2012) used the term African when discussing prototypicality. For consistency, we use the same verbiage here.

voices of Black individuals are centered and uplifted (Black Lives Matter, 2024). #BlackLivesMatter aims to eradicate White supremacy and provide a system of support for Black individuals. Research has begun assessing how individuals perceive the #BlackLivesMatter movement. For instance, recent studies have examined both American women's (Fahs & Swank, 2022) as well as Afro-Latinx's perceptions and support for the movement (Hordge-Freeman & Loblack, 2021). Both studies found that most participants supported the #BlackLivesMatter movement, however to differing degrees (Fahs & Swank, 2022) and rooted in distinct justifications (Hordge-Freeman & Loblack, 2021). Interestingly, a common theme among the #BlackLivesMatter supporters in Hordge-Freeman and Loblack's study (Hordge-Freeman & Loblack, 2021) was that their support was rooted in their awareness of how they are perceived by others. Specifically, participants demonstrated a stark awareness of the reductive nature of perceptions of them by others, such that despite their multifaceted identity, they are often reduced solely to the Black racial component of their identity. Subsequently, they reported experiencing the negative implications associated with Blackness, and therefore, strongly aligned themselves with the #BlackLivesMatter movement.

There is currently a dearth of research examining how evaluator characteristics are associated with supporting the #BlackLivesMatter movement, how people decide who is represented by the movement, and what the implications are of such perceptions. As such, there is still not a clear understanding of how the American public perceives and supports the diverse population of Black Americans that this movement represents. Accordingly, research is needed to investigate how people perceive the diverse range of Black individuals who use the #BlackLivesMatter movement to voice their experiences of race-based hate crimes, especially when Black victims vary in their adherence to expectations surrounding Blackness (e.g., skin tone).

2. Current work

We conducted five experiments, organized as three studies, examining the extent to which people believe Black individuals varying in phenotypicality and prototypicality are represented by the #BlackLivesMatter movement. Notably, we maintained a specific focus on potential differences in perceived hate crime victimization outcomes between targets high and low in phenotypicality. Study 1 examined whether manipulating phenotypicality of Black targets affected evaluations of who is represented by the #BlackLivesMatter movement. Study 2 recruited predominantly White (Study 2a) and predominantly Black (Study 2b) participants to assess how varying phenotypicality impacts perceptions of credibility, trauma, and appropriateness of reporting hate crime victimization on the #BlackLivesMatter website. Study 2c expanded on Studies 2a and 2b by recruiting both White and Black participants to enable the inclusion of participant race as a factor while also explicitly stating target race. Lastly, Study 3 aimed to enhance the ecological validity of our research by assessing hate crime victimization outcomes using mock #BlackLivesMatter hate crime reports. Importantly, Study 3 investigated how a combination of the selected phenotypicality cue (i.e., skin tone) in combination with an external prototypicality cue (i.e., name) influenced perceptions of victim reporting metrics.

3. Study 1

Previous research suggests that Black individuals with more phenotypically Black features tend to experience heightened discrimination and mistreatment (Maddox et al., 2022). However, these phenotypicality biases have yet to be examined within the context of racial social justice movements. Based on these findings, the goal of Study 1 was to test whether phenotypicality biases among Black individuals translate into biases regarding who is represented by the #BlackLivesMatter movement. Skin tone is the predominant feature used when judging

Black racial phenotypicality, but other visible features, including hairstyle and texture, face structure, and dress, can all contribute. To test our hypothesis, participants were asked to indicate the extent to which they would classify 20 targets varying in Black racial phenotypicality (high, low) as represented by the #BlackLivesMatter movement. The resulting categorizations served as our high prototypic and low prototypic clusters.

3.1. Method

Statistical Power and Participants. We determined the sample size for Study 1 based on the sample size used in conceptually similar research (Goh & McCue, 2021). Using this reference point, we targeted approximately 150 participants. Of the 162 participants who responded, 7 did not consent to participate at the beginning of the study and 5 revoked their consent after debrief. As a result, responses from a total of 150 participants ($M_{age} = 39.27$, $SD = 10.78$; 66.0 % men, 32.0 % women, 0.67 % non-binary, 1.3 % preferred not to say; 73.33 % White, 9.33 % Black, 6.67 % East Asian, 3.33 % Bi- or Multiracial, 2.67 % Southeast Asian, 1.33 % Latino/a, 0.67 % South Asian, 0.67 % West Asian, 0.67 % preferred to specify: Caribbean, 1.33 % preferred not to say) were included in our analyses. For all studies, participants were recruited via Connect, an online source of high-quality participants powered by CloudResearch. Given the closely related nature of our studies, each study was run sequentially, and we used the platform's exclusion function to ensure individuals only participated in one study of the current research project. Data for all studies were collected prior to data analysis, and all measures, manipulations, and exclusions are reported in these studies and can be accessed online: https://osf.io/fwq9 z/?view_only=fa92dbea1ea7401bb69170bbc4654453. A sensitivity analysis revealed that when examining the difference between two dependent means, Study 1 provided 80 % power to detect an effect size of $d = 0.21$.

3.1.1. Materials

Stimuli comprised photographs of 20 Black targets differing in Black racial phenotypicality sourced from the Chicago Face Database (CFD; Ma et al., 2015). Based on methods described in Deska et al. (2020), we created low and high Black racial phenotypicality sets using the CFD's norming data and codebook. Specifically, among the images in the CFD labeled as either Black Females ($n = 255$) or Black Males ($n = 253$), the top five pre-rated lowest and highest means for both sexes[2] in terms of Black prototypicality[3] comprised Study 1 stimuli. From the 20 images selected, we ran preliminary analyses to ensure that the images did not significantly differ in terms of attractiveness. Further, we confirmed that prototypicality ratings between high and low conditions did significantly differ. Finally, we confirmed prototypicality ratings between male and female images, but within condition (high, low), did not significantly differ. For a full description of the image selection process, see Supplementary Materials which can be accessed via the study's Open Science Framework page. Therefore, the final stimuli set comprised 10 photos of females (five low, five high) and 10 photos of males (five low, five high). All targets had neutral expressions, were in colour, and configured at 600 width × 400 height. Because the focus of the current work was on racial phenotypicality, and because we did not have specific a priori hypotheses, we did not include analyses across target sex here. Interested readers can find analyses that incorporate target sex in the Supplementary Materials.

---

[2] Note the use of the term "sex" is due to the original labeling of the sourced images from the Chicago Face Database.

[3] The CFD assessed prototypicality through the following operationalization: "How prototypic of the following racial categories would you rate the above target? 1–7; not at all, extremely): White, Black, East Asian, South Asian Indigenous, Latino/x/a."

3.1.2. Procedure

After providing informed consent, participants were provided the following definition of the Black Lives Matter movement: "*The #BlackLivesMatter was founded in 2013 in response to Trayvon Martin's death. Black Lives Matter Global Network Foundation, Inc. is a global organization in the US, UK, and Canada, whose mission is to eradicate White supremacy and build local power to intervene in violence inflicted on Black communities by the state and vigilantes. By combating and countering acts of violence, creating space for Black imagination and innovation, and centering Black joy, #BlackLivesMatter works to make immediate improvements in our lives.*" After viewing the definition, participants saw all 20 photos in a randomized order. For each photograph, participants were asked to indicate how likely they would be to classify the target as represented by the Black Lives Matter movement on a 7-point Likert scale ranging from 1 (*Not at all likely*) to 7 (*Extremely likely*). After completing the prototypicality rating, participants continued to the next trial. Following the completion of all 20 ratings, participants were asked to indicate how familiar they were with the Black Lives Matter Movement on a 7-point Likert scale, ranging from 1 (*Not familiar at all*) to 7 (*Extremely familiar*). Finally, participants completed a demographic survey (e.g., age, racial identity, gender) and were fully debriefed.

3.2. Results

Our primary interest was the extent to which perceptions of representation by the #BlackLivesMatter movement differed between targets low and high in Black racial phenotypicality, irrespective of target sex. As such, we collapsed our four conditions across target sex to create two composite conditions: low and high Black racial phenotypicality.

Following this data re-shaping, and to test this hypothesis, we used paired samples *t*-tests to compare high and low phenotypicality targets on the perceived representativeness measure (see Fig. 1). In line with our hypothesis, participants rated targets high in Black racial phenotypicality ($M = 6.16$, $SD = 1.17$) as more represented by the Black Lives Matter movement than targets low in Black racial phenotypicality ($M = 4.32$, $SD = 1.40$), $t(149) = -14.013$, $p < .001$, 95 % CI $[-2.10, -1.58]$,[4] $d = 1.14$.

We also included a familiarity with the Black Lives Matter movement measure. Overall, there was high familiarity with the Black Lives Matter movement ($M = 5.73$, $SD = 1.23$) and thus we did not conduct further analyses with this measure.
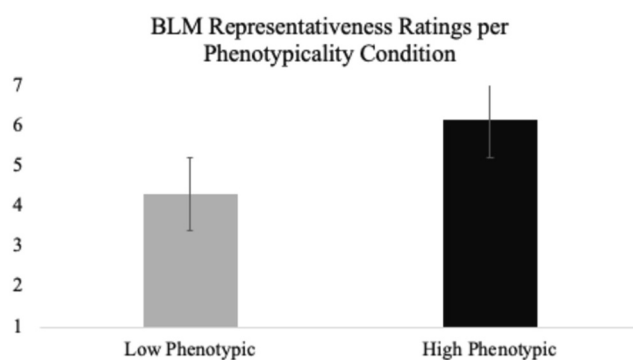

Fig. 1. Perceived representativeness by the Black Lives Matter movement by Black Phenotypicality Ratings in Study 1. Error bars indicate standard error of the mean.

---

[4] Note that the 95 % confidence intervals reported throughout the current manuscript are structured around the mean difference of the sample estimate.

3.3. Discussion

Study 1 assessed the impact of varying phenotypicality level of Black targets and how such factors influence perceptions of who is represented by the #BlackLivesMatter movement. Supporting our primary hypothesis, participants rated targets high in Black phenotypicality as more representative of the #BlackLivesMatter movement compared to targets low in Black phenotypicality. Considering how important physical features, like skin tone, are in racial categorization (Maddox & Gray, 2002), and the nature of #BlackLivesMatter movement as a racial social justice movement, it is perhaps not surprising that darker skinned targets (i.e., high phenotypic condition) were perceived as more representative of this movement compared to their lighter skinned (i.e., low phenotypic condition) counterparts. However, it is important to note that although skin tone is primary in judging racial prototypicality (Maddox, 2004), it is not the only feature that contributes to these judgments, and the current work is not designed to test the relative impact of these features. Together, these initial findings are consistent with our primary hypothesis regarding the impact of phenotypicality on perceptions of representation within the #BlackLivesMatter movement.

4. Study 2a

Study 1 indicated that participants believe that targets low in Black racial phenotypicality are less represented by the Black Lives Matter movement compared to targets high in Black racial phenotypicality. In Study 2a, we investigated several implications for these prototypicality judgments. Specifically, we tested the extent to which judgments of credibility, harm, and appropriateness of reporting on the #BlackLivesMatter website following claims of hate crime victimization were influenced by target racial phenotypicality.

In Study 2a, participants viewed four images of targets that differed in Black racial phenotypicality (i.e., low, high) and sex (i.e., male, female), and viewed a brief written vignette that read: "the individual depicted above has claimed that they are the victim of a hate crime and is planning to report the incident on the Black Lives Matter online-reporting system." Using similar procedures outlined in previous research (Goh et al., 2021), we then asked participants to indicate the extent to which they believed the target was the victim of a hate crime, how traumatized they think they would be, and how appropriate they believe it is that the target is choosing to report the crime on the #BlackLivesMatter website. We hypothesized that participants would perceive high racially phenotypic targets as more credible, more traumatized, and that it was more appropriate for them to use the #BlackLivesMatter website compared to low racially phenotypic targets.

4.1. Method

Statistical power and participants. We determined the sample size for Study 2a based on the sample size used in a conceptually similar study in the literature (Goh & McCue, 2021). Using this reference point, we targeted approximately 150 participants. All participants over the age of 18 years old who lived in the United States were eligible to participate, and no participants were excluded from analyses.

A total of 150 Cloud Research workers ($M_{age} = 38.45$, $SD = 10.11$; 60.0 % men, 37.33 % women, 2.0 % non-binary, 0.67 % preferred to specify; 70.0 % White, 9.33 % Black, 8.67 % Latino/a, 4.67 % East Asian, 3.33 % Biracial, 2.00 % preferred not to say, 1.33 % South Asian, 0.67 % Southeast Asian) participated in the study. A sensitivity analysis revealed that when examining the difference between two dependent means, Study 2a provided 80 % power to detect an effect size of $d = 0.21$.

Materials. Because we added several measures and a vignette, we showed participants fewer targets to reduce fatigue. For each phenotypicality × sex combination, we selected targets with the highest average prototypicality score from Study 1 (see Table 1). Consequently, stimuli

Table 1
Prototypicality means by condition and image from Study 1.

Image #	Condition						
	Male				Female		
	Low		High		Low		High
	M SD		*M SD*		*M SD*		*M SD*
1	4.03 (1.86)		6.13 (1.33)		3.57 (2.01)		6.17 (1.34)
2	3.95 (1.81)		6.10 (1.28)		4.08 (1.80)		6.18 (1.35)
3	4.45 (1.67)		6.29 (1.22)		4.33 (1.81)		6.16 (1.31)
4	4.20 (1.81)		6.17 (1.23)		5.36 (1.45)		6.21 (1.31)
5	4.31 (1.84)		6.12 (1.33)		4.91 (1.71)		6.11 (1.24)

for Study 2a comprised four photographs (two low phenotypic, two high phenotypic targets; two females, two males).

Procedure. We provided participants with the same definition of the #BlackLivesMatter movement provided in Study 1. After viewing the definition, participants saw all four photos in a randomized order. For each photograph, participants read a brief written vignette describing that the depicted target claimed they were the victim of a hate crime and is planning to report the incident on the #BlackLivesMatter online-reporting system. The following definition of a hate crime was provided: "*A crime motivated by bias against race, colour, religion, national origin, sexual orientation, gender, gender identity, or disability*" (U.S. Department of Justice, 2023). After viewing the image of the target as well as the written vignette, participants indicated how much that they believe the target was the victim of a hate crime (credibility; 1 = *Do not believe at all*, 7 = *Definitely believe*), how traumatized they believe the individual would be if they were a victim of a hate crime (harm; 1 = *Not traumatized at all*, 7 = *Extremely traumatized*), and how appropriate they believe it is that the individual is reporting the incident on the #BlackLivesMatter online reporting system (appropriateness; 1 = *Not at all appropriate*, 7 = *Extremely appropriate*). After completing all ratings, participants continued to the next trial. Following the completion of all trials, participants were asked to indicate how familiar they were with the #BlackLivesMatter movement using the same measure as Study 1. Finally, participants completed a demographic survey (e.g., age, racial identity, gender) and were fully debriefed.

4.2. Results

Similar to our primary hypothesis for Study 1, the primary goal of Study 2a was to examine perceived differences by Black racial phenotypicality level. As such, we first averaged participants' perceived victim credibility, harm, and appropriateness ratings for the low and the high phenotypicality targets to create composite scores for each group. As our primary interest was the extent to which phenotypicality influences implications of hate crime victimization (i.e., perceived victim credibility and harm, victim reporting), we analyzed results using three paired samples *t*-tests.

Results of these analyses supported two of our three primary hypotheses (see Fig. 2). Specifically, participants rated high phenotypic targets ($M = 5.13$, $SD = 1.75$) as more credible than low phenotypic targets ($M = 4.86$, $SD = 1.74$), $t(149) = -3.47$, $p < .001$, $d = 0.28$, 95 % CI [-0.42, -0.17]. Further, it was deemed more appropriate for high phenotypic targets ($M = 5.21$, $SD = 1.97$) to report the incident on the #BlackLivesMatter website than low phenotypic targets ($M = 4.86$, $SD = 2.06$), $t(149) = -3.75$, $p < .001$, $d = 0.31$, 95 % CI [-0.53, -0.16]. However, the results did not support our hypothesis related to differences in perceptions of harm following hate crime victimization. No significant differences emerged between low phenotypic targets ($M = 5.39$, $SD = 1.55$) and high phenotypic targets ($M = 5.47$, $SD = 1.63$) in perceptions of harm following victimization, $t(149) = -1.30$, $p = .1953$, $d = 0.11$, 95 % CI [-0.21, 0.04]. Results of correlational analyses revealed that credibility, harm, and appropriateness were all significantly, positively correlated with each other (see Supplementary Materials).

Identical to the rationale presented in Study 1, given the relatively high familiarity with the Black Lives Matter movement reported by participants ($M = 5.64$, $SD = 1.23$), no further analyses using this variable were conducted.

4.3. Discussion

In Study 2a, we assessed how Black racial phenotypicality impacted perceived credibility, harm, and appropriateness of reporting a hate crime to the #BlackLivesMatter website. Our findings supported two of our three hypotheses. Specifically, high phenotypic targets were rated as more credible and participants deemed that it was more appropriate for them to report using the movement's website compared to low phenotypic targets. However, in contrast to our hypothesis, there were no significant differences found between low and high phenotypic targets regarding perceived harm of the hate crime incident. A similar study by Pejic and Deska (Study 2; 2024) also found no differences in perceived harm following an anti-Asian hate crime between low and high Asian phenotypic targets.

The majority of participants in Study 2a were White, and research has demonstrated White evaluators are less sensitive to skin tone biases of intergroup members compared to intragroup members (Maddox, 2004). Further, as perceptions of Black racial phenotypicality have been assessed using both Black and White samples (e.g., Maddox & Gray, 2002), we similarly wanted to assess whether there were any differences in perceptions of hate crime outcomes using a primarily Black sample. As such, the primary aim of Study 2b was to assess how Black racial phenotypicality influences perceptions of race-based hate crime victimization outcomes with a primarily Black sample.

5. Study 2b

Study 2a demonstrated that among predominantly White participants, Black racial phenotypicality influenced perceptions of credibility and appropriateness of reporting hate crime victimization on the #BlackLivesMatter website. Notably, no differences in perceived harm experienced following hate crime victimization were observed between targets low and high in Black racial phenotypicality.

For Study 2b, we were similarly interested in outcomes associated with prototypicality judgments following hate crime victimization. However, this study differed in that we recruited predominantly Black participants. Specifically, we sought to test how Black racial phenotypicality influences perceptions of hate crime victimization among a predominantly Black sample.

5.1. Method

Statistical Power and Participants. As the goal of Study 2b was to replicate Study 2a, with the notable exception of additional inclusion criteria, the equivalent sample size of 150 participants was targeted. All participants over the age of 18 years old who lived in the United States, and self-identified as Black were eligible to participate, and no participants were excluded from analyses.

A total of 150 Cloud Research workers ($M_{age} = 35.81$, $SD = 9.01$; 49.33 % women, 47.33 % men, 2.67 % non-binary, 0.67 % preferred to

Mean Ratings of Outcome Variables per Phenotypicality Condition
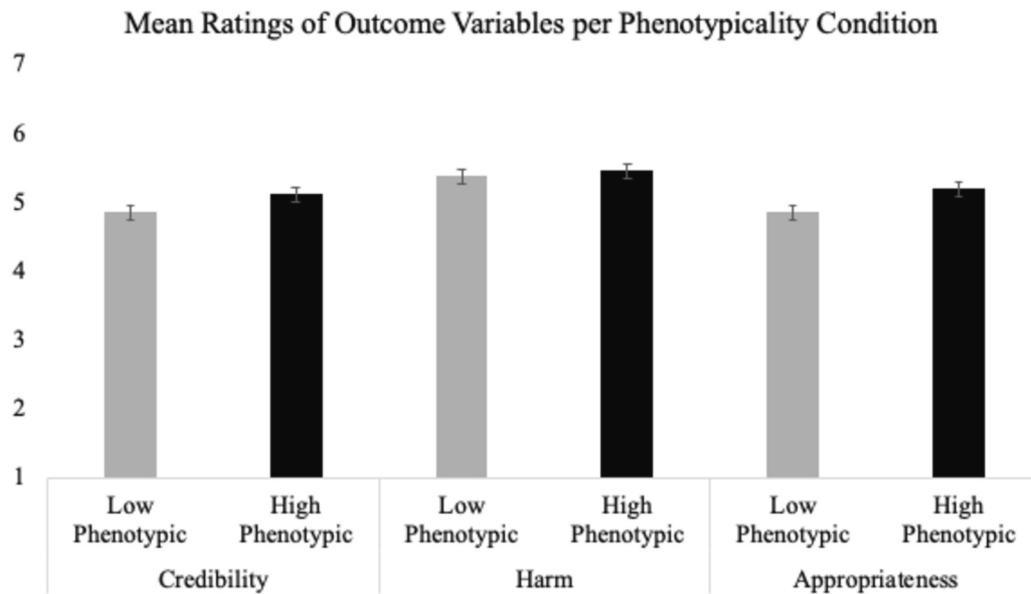

Fig. 2. Mean ratings by condition in Study 2a. Error bars indicate standard error of the mean.

specify; 92.67 % Black, 3.33 % Bi-Racial, 1.33 % Latino,[5] 1.33 % White, 0.67 % preferred to specify, 0.67 % preferred not to specify) participated in the study. A sensitivity analysis revealed that when examining the difference between two dependent means, Study 2b provided 80 % power to detect an effect size of $d = 0.21$.

5.1.1. Materials

The stimuli used for Study 2b was identical to the stimuli used for Study 2a.

5.1.2. Procedure

The procedure for Study 2b was identical to the procedure for Study 2a.

5.2. Results

Expanding on Study 2a, our primary interest was the extent to which Black racial phenotypicality influences implications of hate crime victimization among primarily Black participants. As such, three paired samples *t*-tests were performed to compare the effect of phenotypicality (low, high) on our three hate crime victimization outcome variables (credibility, harm, appropriateness).

Identical to Study 2a's analysis plan, we first averaged participants' perceived victim credibility, harm, and appropriateness ratings to create composite scores for each high and low phenotypicality groups, and then submitted these values to a series of paired samples t-tests.

Results of these analyses partially replicated findings from Study 2a (see Fig. 3). Specifically, among primarily Black participants, targets low in Black racial phenotypicality were similarly rated as less credible ($M = 5.17$, $SD = 1.32$), and it was deemed less appropriate for them to report their hate crime victimization on the #BlackLivesMatter reporting site

---

[5] Despite the eligibility criteria for participation in the study, a small percentage of non-Black individuals participated in the study. Given the limited ineligible individuals who participated, in combination with the desire to reduce unnecessary researcher degrees of freedom, we included all participants' data in the current paper. As such, we refer to the current sample as, "primarily Black participants." However, all analyses were run with only Black participants for comprehensiveness purposes, and the pattern of findings did not change.
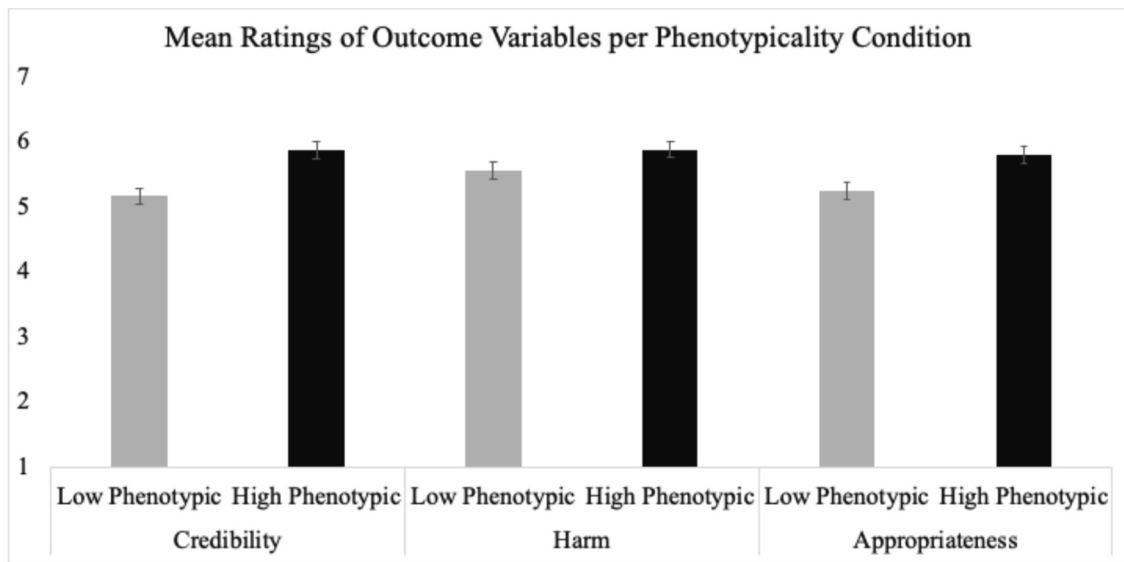
($M = 5.25$, $SD = 1.51$), compared to their high prototypic counterparts for both credibility ($M = 5.87$, $SD = 1.15$), $t(149) = -7.90$, $p < .001$, $d = 0.65$, 95 %CI[-0.88, -0.53], and appropriateness ($M = 5.81$, $SD = 1.53$), $t(149) = -7.69$, $p < .001$, $d = 0.63$, 95 %CI[-0.70, -0.41].

However, when analyzing differences in perceptions of harm between low and high Black phenotypic targets, results were in line with our original hypothesis, but not in line with Study 2a's findings. Specifically, targets low in Black racial phenotypicality were rated as significantly less harmed ($M = 5.57$, $SD = 1.27$) following alleged hate crime victimization compared to their high racially phenotypic counterparts ($M = 5.89$, $SD = 1.08$), $t(149) = -4.11$, $p < .001$, $d = 0.34$, 95 % CI[-0.48, -0.17]. Consistent with Study 1 and 2a, participants reported high levels of familiarity with the Black Lives Matter movement ($M = 5.93$, $SD = 1.12$). Results of correlational analyses revealed that credibility, harm, and appropriateness were all significantly, positively correlated with each other (see Supplementary Materials).

5.3. Discussion

Consistent with our hypotheses, our primarily Black sample perceived targets low in Black racial phenotypicality as less credible, less appropriate to report their hate crime incident on the #BlackLivesMatter website, and that they were less harmed compared to targets high in Black racial phenotypicality. The current research allowed us to more thoroughly assess perceptions of hate crime victimization by consecutively assessing both outgroup and ingroup evaluations. Whereas our primarily Black sample in the current study rated low Black phenotypic targets as enduring less harm following the incident compared to high Black phenotypic targets, the majority White sample in Study 2a did not perceive differences in harm endured. Similarly, Strom et al. (2012) found that White perceivers were less sensitive to skin tone variation when evaluating the prototypicality of Black targets compared to Black perceivers, as skin tone was an integral cue for Black perceivers when assessing prototypicality of Black targets. Thus, perhaps Black perceivers are not only highly sensitive to appearance cues such as skin tone variation when assessing Black targets' prototypicality, but perhaps such evaluations influence their assessment of outcomes associated with Black targets, such as perceived harm endured. Further, as Black individuals will have a unique perspective based on their lived experiences with hate crime victimization and that of close others (e.g., family, friends), their personal experiences with

Fig. 3. Mean ratings by condition in Study 2b. Error bars indicate standard error of the mean.

hate crime victimization may further influence these judgments.

One notable limitation of Studies 1-2b is that we selected images based on racial prototypic ratings provided by the CFD's norming data. These values were calculated by the number of participants who categorized the model as Black divided by the number of participants who rated the model (i.e., the proportion of people who correctly identified the model as Black). Thus, although all the selected images were of individuals who self-identify as Black, not all raters categorized them as such (e.g., high prototypical targets had a higher likelihood of being classified as Black compared to lower prototypic targets). Therefore, it is possible that racial categorization served as a confound in these studies, and that the findings may have been caused by low-prototypic targets having their race misclassified. We sought to address this limitation in Study 2c.

6. Study 2c

Study 2a and 2b examined whether Black racial phenotypicality influenced perceptions of credibility, harm, and appropriateness of reporting hate crime victimization on the #BlackLivesMatter website among both Black and White participants. Although images included in the low and high phenotypic conditions within Study 2a and 2b were independently rated as Black before the completion of each study, absent from both study's vignettes was an explicit statement indicating that the target depicted in the image is Black. As such, to ameliorate concerns regarding the believability about targets' race (e.g., perceptions of non-Black ethnicity or bi-racial), Study 2c aimed to re-test our hypotheses with an amended vignette. We recruited both Black and White participants to allow the inclusion of participant race as a factor in our analyses.

6.1. Method

Statistical Power and Participants. As the goal of Study 2c was to replicate Study 2a and 2b, with the notable exception of an amended vignette, the equivalent sample size of 300 participants (150 White, 150 Black) was targeted. All participants over the age of 18 years old who lived in the United States, and self-identified as Black or White were eligible to participate, and no participants were excluded from analyses.

A total of 276 Cloud Research workers ($M_{age} = 37.30$, $SD = 11.90$; 48.90 % women, 49.60 % men, 1.1 % non-binary, 0.40 % preferred to specify; 48.90 % Black, 51.10 % White) participated in the study. A

sensitivity analysis revealed that when examining the difference between two dependent means, Study 2c provided 80 % power to detect an effect size of $d = 0.14$.

6.1.1. Materials

The visual stimuli used for Study 2c was identical to the stimuli used for Study 2a and 2b. However, the amended vignette read as follows: *"The individual depicted above is a Black [man/woman] and has claimed that they are the victim of a hate crime and is planning to report the incident on the Black Lives Matter online-reporting system."*

6.1.2. Procedure

The procedure for Study 2c was identical to the procedure for Study 2a and 2b.

6.2. Results

Expanding on Study 2a and 2b, our primary interest was the extent to which Black racial phenotypicality influences implications of hate crime victimization among primarily Black and White participants when the race of the target was explicitly stated. As such, three repeated measures ANOVAs were performed to compare the effect of phenotypicality (low, high) on our three hate crime victimization outcome variables (credibility, harm, appropriateness) between both White and Black participants.

Identical to Study 2a and 2b's analysis plan, we first averaged participants' perceived victim credibility, harm, and appropriateness ratings to create composite scores for each high and low phenotypicality groups, and then submitted these values to a series of repeated measures ANOVAs.

To first test whether phenotypicality (low, high) influenced judgments of credibility following hate crime victimization among Black and White participants, we submitted the averaged credibility ratings to a 2 (Participant Race: White, Black) X 2(Phenotypicality Status: High, Low) repeated measures ANOVA. Results revealed a main effect of phenotypicality status on credibility ratings, $F(1, 274) = 38.40$, $p < .001$, $\eta_p^2 = 0.12$. Consistent with the findings of 2a and 2b, targets high in Black racial phenotypicality ($M = 5.49$, $SD = 1.51$) were rated as more credible following hate crime victimization compared to targets low in Black racial phenotypicality ($M = 5.17$, $SD = 1.49$). There was a main effect of participant race on credibility ratings, $F(1, 274) = 24.40$, $p < .001$, $\eta_p^2 = 0.08$. Specifically, Black participants ($M = 5.75$, $SD = 1.21$) perceived

the high phenotypic and low phenotypic targets to be more credible than White participants ($M = 4.93$, $SD = 1.64$) did across the same high and low phenotypic conditions. Further, results revealed a significant interaction between participant race and phenotypic status, $F(1, 274) = 15.10$, $p < .001$, $\eta_p^2 = 0.05$ (see Fig. 4).

Decomposing this interaction revealed Black participants perceived high phenotypic targets ($M = 6.01$, $SD = 1.12$) as more credible compared to low phenotypic targets ($M = 5.48$, $SD = 1.30$), $t(274) = 7.05$, $p < .001$, $d = 0.43$. However, there was no significant difference in credibility ratings between high ($M = 4.99$, $SD = 1.67$) and low phenotypic targets ($M = 4.87$, $SD = 1.60$) among White participants, $t(274) = 1.65$, $p = .351$, $d = 0.01$. Comparing across participant race revealed that Black participants ($M = 6.01$, $SD = 1.12$) perceived high phenotypic targets as more credible compared to White participants ($M = 4.99$, $SD = 1.67$), $t(274) = 5.96$, $p < .001$, $d = 0.71$. Finally, Black participants ($M = 5.48$, $SD = 1.30$) perceived low phenotypic targets to be significantly more credible compared to White participants ($M = 4.87$, $SD = 1.60$), $t(274) = 3.50$, $p = .003$, $d = 0.42$.

To test whether phenotypicality (low, high) influenced judgments of appropriateness of reporting following hate crime victimization among Black and White participants, we submitted the averaged appropriateness ratings to a 2(Participant Race: White, Black) X 2(Phenotypicality Status: High, Low) repeated measures ANOVA. Results revealed a significant main effect of phenotypicality on appropriateness ratings, $F(1, 274) = 23.87$, $p < .001$, $\eta_p^2 = 0.08$. Replicating our previous findings, participants rated it more appropriate for high phenotypic targets ($M = 5.53$, $SD = 1.76$) to report their victimization on the BLM website compared to low phenotypic targets ($M = 5.31$, $SD = 1.75$). There was also a main effect of participant race on appropriateness ratings, $F(1, 274) = 14.40$, $p < .001$, $\eta_p^2 = 0.05$. Black participants ($M = 5.81$, $SD = 1.34$) perceived the targets as more appropriate in their reporting compared to White participants ($M = 5.05$, $SD = 2.01$) ratings of the same targets. Finally, there was a significant interaction between participant race and phenotypicality status on appropriateness ratings, $F(1, 274) = 5.60$, $p = .019$, $\eta_p^2 = 0.02$ (see Fig. 5).

Decomposing this interaction revealed that Black participants deemed it more appropriate for high phenotypic targets ($M = 5.98$, $SD = 1.29$) to report their victimization compared to low phenotypic targets ($M = 5.64$, $SD = 1.38$), $t(274) = 5.07$, $p < .001$, $d = 0.31$. However, there was no significant difference in appropriateness ratings between high

phenotypic targets ($M = 5.11$, $SD = 2.03$) and low phenotypic targets ($M = 4.99$, $SD = 1.99$) among White participants, $t(274) = 1.80$, $p = .275$, $d = 0.11$. Comparing across participant race revealed that Black participants ($M = 5.98$, $SD = 1.29$) deemed it was more appropriate for high phenotypic targets to report their victimization compared to White participants ($M = 5.11$, $SD = 2.03$), $t(274) = 4.24$, $p < .001$, $d = 0.51$. Finally, Black participants ($M = 5.64$, $SD = 1.38$) rated low phenotypic targets as significantly more appropriate to report their victimization compared to White participants ($M = 4.99$, $SD = 1.99$), $t(274) = 3.15$, $p = .010$, $d = 0.38$ (See Fig. 6).

Finally, to test whether phenotypicality (low, high) influenced judgments of harm following hate crime victimization among Black and White participants, we submitted the averaged harm ratings to a final 2 (Participant Race: White, Black) X 2(Phenotypicality Status: High, Low) repeated measures ANOVA. Results revealed a significant main effect of phenotypicality level on harm ratings, $F(1, 274) = 3.96$, $p = .047$, $\eta_p^2 = 0.01$. Specifically, participants rated targets in the high phenotypic condition ($M = 5.79$, $SD = 1.40$) as more harmed following their hate crime victimization compared to targets in the low phenotypic condition ($M = 5.70$, $SD = 1.33$). Results also supported a main effect of participant race, $F(1, 274) = 10.90$, $p = .001$, $\eta_p^2 = 0.04$. Specifically, Black participants ($M = 6.01$, $SD = 1.16$) perceived targets in both the high and low phenotypic conditions as more harmed compared to White participants ($M = 5.50$ $SD = 1.50$) perceptions of harm in the high and low phenotypic conditions. Participant race and target phenotype status did not significantly interact, $F(1, 274) = 0.95$, $p = .332$, $\eta_p^2 = 0.00$. Results of correlational analyses revealed that credibility, harm, and appropriateness were all significantly, positively correlated with each other (see Supplementary Materials).

6.3. Discussion

Study 2c replicated the results of Study 2b while explicitly labelling targets as Black. As main effects, we again observed that participants viewed targets low in Black racial phenotypicality as less credible following hate crime victimization, judged it was less appropriate for low phenotypic targets to report their crime on the #BlackLivesMatter website, and rated high phenotypic targets as more harmed following their hate crime victimization compared to low phenotypic targets. Participant race interacted with target phenotype status on ratings of
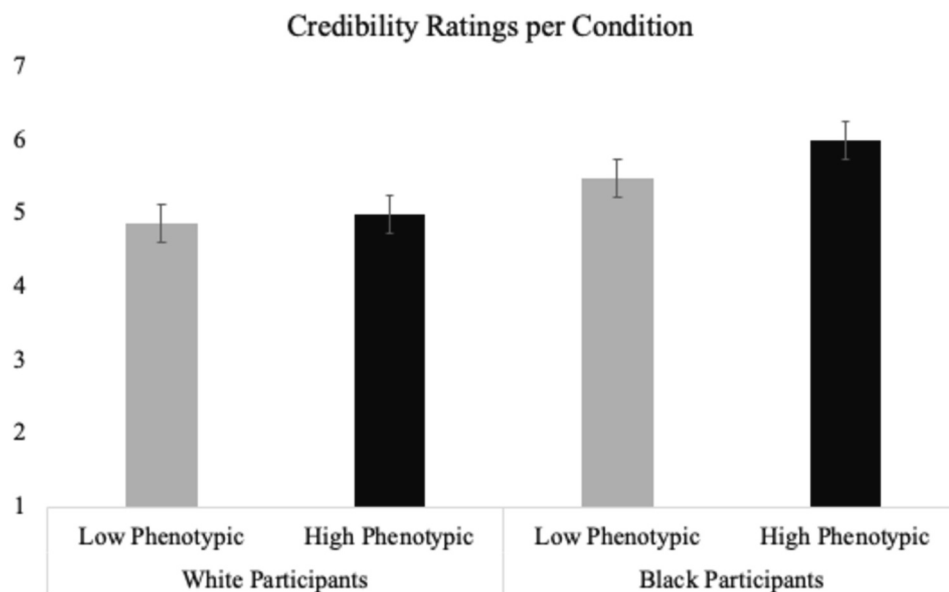
Fig. 4. Mean credibility ratings by condition in Study 2c. Error bars indicate standard error of the mean.
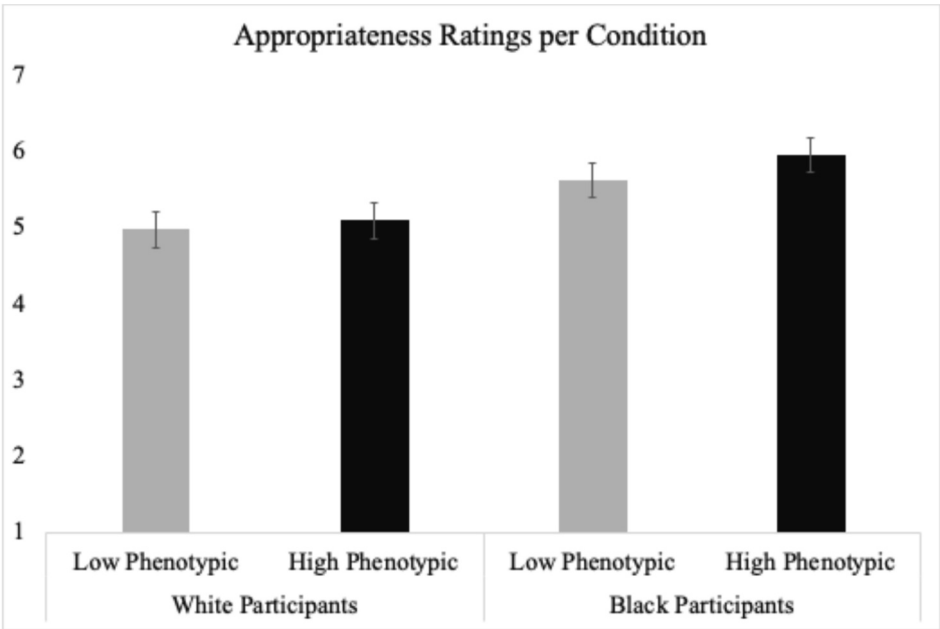
Fig. 5. Mean appropriateness ratings by condition in Study 2c. Error bars indicate standard error of the mean.
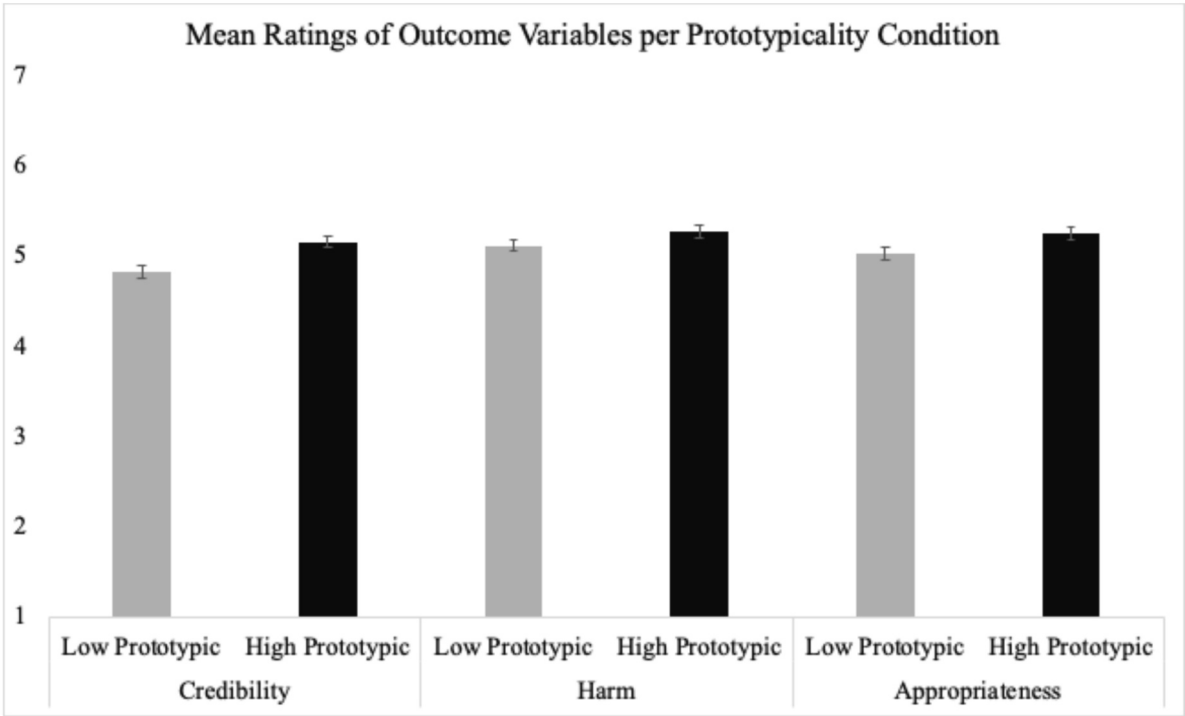

Fig. 6. Mean ratings by condition in Study 3. Error bars indicate standard error of the mean.

credibility and appropriateness, but not harm. Most notably, these interactions revealed that the main effects of credibility and appropriateness were both driven by Black participants, and that White participants did not differentiate between high and low phenotypic targets when rating credibility or appropriateness. There were also main effects of participant race for credibility, appropriateness, and harm such that Black participants viewed both the high and low phenotypic targets as more credible, more appropriate in their reporting, and as having experienced more harm than did White participants ratings of the same high and low targets. These discrepancies across participant race may speak to broader issues about how prejudice and

discrimination are perceived in our communities. It may be the case that Black participants are more sensitive to hate crimes in general than White participants due to greater collective trauma, experiences with hate crimes, and empathy for victims (e.g., Bryant-Davis et al., 2017). Related, White participants may lack sensitivity to hate crimes to justify their worldviews (Jost & Banaji, 1994), maintain the social hierarchy (Pratto et al., 1994), or due to a relative inability to simulate the experience. Indirect evidence for this comes from recent work showing that, after fatal police shootings, White participants tend to sympathize with police officers compared to racial minority victims (Kahn et al., 2024). Ultimately, more research is needed to better understand how

various participant factors influence perceptions of hate crime reporting metrics.

7. Study 3

Studies 2a-c provided evidence for potential downstream implications of the observed prototypicality bias among White and Black samples. Specifically, these studies demonstrated that when a target is depicted as being the victim of a hate crime, targets' high in Black racial phenotypicality claims were viewed as more credible, and it was deemed more appropriate for them to use the #BlackLivesMatter online reporting system compared to their low Black phenotypic counterparts. Notably, these biases were generally demonstrated among both Black and White participants, albeit with some variation across samples in Studies 2a-c.

In Study 3, we sought to conceptually replicate our previous results with a different manipulation of prototypicality and in a more ecologically valid context. Specifically, Study 3 expanded on our previous design by recruiting participants for a study evaluating the clarity of victim reports from sociopolitical movements' reporting platforms. Participants were told they were randomly assigned to view hate crime reports, then were presented with four fictional hate crime reports in randomized order, and informed that they were generated from the #BlackLivesMatter reporting site. Notably, this final study differed in that images of targets were no longer provided to participants. Rather, Black racial prototypicality was manipulated via text content within the vignettes (i.e., target name, skin tone). Notably, this study was not designed to test the relative importance of name and skin tone as sources of prototypicality, but rather was designed to make prototypicality salient via these multiple cues.

7.1. Method

Statistical Power and Participants. Identical to the rationale used for Study 2b's target sample, as the goal of Study 3 was to conceptually replicate our previous findings, the equivalent sample size of 150 participants was targeted. All participants over the age of 18 years old who lived in the United States were eligible to participate, and no participants were excluded from analyses.

A total of 150 Cloud Research workers (M_{age} = 39.22, SD = 11.66; 40 % women, 57.33 % men, 2 % non-binary, 67 % preferred to specify; 74 % White, 7.3 % Latino, 6.67 % East Asian, 4 % Black, 2.67 % Bi-Racial, 2 % South Asian, 1.33 % Southeast Asian, 1.33 % Indigenous, 0.67 %, preferred to specify) participated in the study. A sensitivity analysis revealed that when examining the difference between two dependent means, Study 3 provided 80 % power to detect an effect size of d = 0.21.

7.1.1. Materials

Stimuli for Study 3 comprised four crime reports appearing to be generated by the #BlackLivesMatter online-reporting system. Within each report, the willingness to report an act of hate, the date of the hate crime, the specific part of their identity they believed was targeted (race), their race (African American),[6] their physical appearance (light skinned vs. dark skinned), their name (racially prototypic vs. non-

racially prototypic)[7] and their gender (man, woman) was depicted. A description of the exact hate crime was not included in the report. As such, participants were simply instructed that each target was claiming that they were the victim of a hate crime - a crime motivated by an individual's identity, such as their race. Skin tone (i.e., physical appearance description) and names were manipulated to represent targets either low or high in Black racial prototypicality. Jada and Lemarcus were the high prototypic names, and James and Emily were the low prototypic names. The high prototypic names were only paired with the dark skinned victims and the low prototypic names were only paired with the light skinned victims (i.e., this was not a crossed design examining cue type). Aside from descriptions of the targets' physical appearance, name, and gender, all other aspects of the crime report remained constant across conditions.

7.1.2. Procedure

Participants were recruited for a study evaluating how laypersons assess crime reports. After providing informed consent, participants were instructed that they are evaluating how clear reports generated from socio-political movement reporting platforms are to assess. They were then instructed that they were randomly assigned to the #BlackLivesMatter reporting site, and provided the same definition of #BlackLivesMatter as Study 1. Following this definition, participants were then instructed that they will view four, unrelated #BlackLivesMatter crime reports, and that they have been further randomly assigned to view "hate crime reports" to ameliorate any concerns about why a similar crime was committed across all "unrelated" reports. Participants were then provided the same definition of a hate crime as presented in Study 2a, 2b, and 2c. Following this final definition, participants viewed four crime reports in randomized order.

For each crime report, participants indicated how much that they believe the target was the victim of a hate crime (credibility; 1 = *Do not believe at all*, 7 = *Definitely believe*), how traumatized they believe the individual would be if they were a victim of a hate crime (harm; 1 = *Not traumatized at all*, 7 = *Extremely traumatized*), and how appropriate they believe it is that the individual is reporting the incident on the #BlackLivesMatter online reporting system (appropriateness; 1 = *Not at all appropriate*, 7 = *Extremely appropriate*). After completing all ratings, participants continued to the next trial. Following the completion of all trials, participants were asked to indicate how familiar they were with the #BlackLivesMatter movement using the same measure as Study 1. Finally, participants completed a demographic survey (e.g., age, racial identity, gender) and were fully debriefed.

7.2. Results

Identical to the analytic steps for Studies 2a-c, we first averaged participants' perceived victim credibility, harm, and appropriateness ratings across high and low prototypicality vignettes. As our primary interest was the extent to which prototypicality (as operationalized by skin tone phenotypicality cues and prototypical name manipulations) influences implications of hate crime victimization (i.e., perceived victim credibility and harm, victim reporting), we analyzed results using three paired samples *t*-tests.

Results of these analyses replicated our earlier findings (see Fig. 4). Specifically, when presented with crime reports depicting targets varying in racial prototypicality, targets low in Black racial prototypicality were rated as less credible (M = 4.84, SD = 1.48), and it was deemed less

---

[6] Although we recognize that African American is more appropriately considered an ethnic group, we used it in the context of race on this vignette because it mirrors how it is used in US census data and in common parlance.

[7] To determine racially prototypic names, a pilot study was conducted using a convenience sample where participants were asked to rate how "stereotypical" they believed various names were of each five ethnic/racial groups, including Black. Names rated to be most stereotypical of the Black category were used for the "high prototypic" names. Names rated to be the most stereotypical of the White category were used for the "low prototypic" names.

appropriate for them to report their hate crime victimization on the Black Lives Matter reporting site ($M = 5.04$, $SD = 1.77$), compared to their high prototypic counterparts for both credibility ($M = 5.17$, $SD = 1.46$), $t(149) = -4.73$, $p < .001$, 95 %CI[-0.46, -0.19], $d = 0.39$, and appropriateness ($M = 5.26$, $SD = 1.74$), $t(149) = -4.15$, $p < .001$, 95 % CI[-0.34, -0.12], $d = 0.34$. Further, results of the harm analyses replicated those found in Study 2b, but not 2a. Specifically, targets depicted as low in Black racial prototypicality were perceived to be significantly less harmed ($M = 5.13$, $SD = 1.43$) as a result of their alleged hate crime victimization compared to their high prototypic counterparts ($M = 5.29$, $SD = 1.44$), $t(149) = -2.60$, $p = .010$, $d = 0.21$, 95 %CI[-0.28, -0.04]. Consistent across all studies, participants reported high levels of familiarity with the #BlackLivesMatter movement ($M = 5.61$, $SD = 0.98$). Results of correlational analyses revealed that credibility, harm, and appropriateness were all significantly, positively correlated with each other (see Supplementary Materials).

7.3. Discussion

Study 3 extended on our prior findings by conceptually replicating the previous findings through manipulating Black prototypicality with written descriptions instead of visual images of targets, and by using a more ecologically valid context. With a primarily White sample, results of Study 3 conceptually replicated those of Study 2b and 2c. Specifically, low Black prototypic targets were rated as less credible, less harmed, and that it was less appropriate for them to report a hate crime incident using the #BlackLivesMatter website compared to their high prototypic counterparts. Our findings for the current study extend the phenotypicality and prototypicality literature by assessing perceptions of outcomes associated with Black racial prototypicality using written prototypicality and phenotypicality cues (e.g., target name, skin tone). There may be some contexts in which written descriptions of hate crime victims are be evaluated without accompanying images (e.g., #BlackLivesMatter report, police report), and thus study provides evidence that the observed phenotypicality and protoypicality bias extends beyond visual cues to include written descriptions as well. One interesting aspect of this study is that, like Study 2a, our sample was primarily White. However, unlike Study 2a, we did find an effect of prototypicality on evaluations of harm endured, mirroring Study 2b and 2c. One possibility is that the observed bias may be stronger for White perceivers with written cues compared to visual cues. One limitation of this study is that, although we provided a definition of hate crime, we did not provide any details regarding the type of hate crime that occurred. This omission may have resulted in participants differentially interpreting the hate crime event across targets and/or conditions. However, we suspect that the effect may simply have been stronger if eliminated these extra potential sources of variance by holding hate crime constant. Regardless, future replications should examine and manipulate the type of hate crime included in the vignettes.

8. General discussion

Phenotypicality cues, such as skin tone, tend to guide our categorization and understanding of individuals' racial identities. Our evaluations of Black individuals are no exception to this, as Black individuals with darker skin tone tend to be perceived as more prototypical of the Black racial category than their lighter skin counterparts (Maddox, 2004). Considering the racial and ethnic diversity among Black Americans, the reductionist outcomes associated with relying on external physical characteristics as cues for racial categorization evaluations fails to adequately capture the diversity in this group. The #BlackLivesMatter movement aims to center Black voices and celebrate the diversity within the Black community. Although some research has examined perceptions of the #BlackLivesMatter movement in relation to evaluators' identity components (e.g., evaluators' gender and race; Fahs & Swank, 2022; Hordge-Freeman & Loblack, 2021), limited research has

examined whether evaluators perceive all Black Americans to be equally represented by the social justice movement as intended. Further, no identifiable research has examined the potential negative consequences of narrow prototypes regarding social justice movement inclusion, such as differential perspectives on hate crime reporting metrics via a racially-specific outlet. As such, the current research aimed to fill in this gap in the literature regarding how Black racial phenotypicality shapes evaluations of race-based hate crime victimization outcomes, with a specific focus examining how Black racial phenotypicality impacts perceived representation within the #BlackLivesMatter movement.

Results of Study 1 supported and extended prior literature by indicating that high phenotypic Black individuals are perceived as being more representative of a social justice movement, #BlackLivesMatter, compared to low phenotypic Black individuals. As darker skinned Black individuals tend to be perceived as more representative of the Black racial category (i.e., more prototypical; Maddox, 2004), our research extends the prototypicality literature to the social justice domain. Studies 2a-c expanded on the implications associated with perceived representation of the #BlackLivesMatter movement by assessing race-based hate crime victimization outcomes of victim credibility, trauma, and appropriateness of reporting the incident on the #BlackLivesMatter website. We tested our hypotheses using both a predominantly White (Study 2a), predominantly Black (Study 2b), and Black and White (Study 2c) sample address intra- and intergroup perspectives. Study 2c also explicitly labeled target race. Although there was some variability around which participant showed precisely which effects across studies (e.g., only Black participants showed an effect on harm), we generally observed effects such that participants viewed high phenotypic Black targets as being more credible, more harmed, and that it was more appropriate for them to report the hate crime incident on the #BlackLivesMatter website than the low phenotypic Black target. Study 3 conceptually replicate the findings of Study 2a-c by assessing hate crime victimization outcomes associated with Black prototypicality using a different manipulation that used both phenotypicality (i.e., skin tone) and prototypicality cues (i.e., name) in written descriptions, and in a more ecologically valid context, assessing #BlackLivesMatter hate crime reports. Results indicated that high Black prototypic targets were rated as more credible, traumatized, and that it was more appropriate for them to report their hate crime incident to the #BlackLivesMatter website compared to low prototypic Black targets. Interestingly, Study 3's predominantly White sample did show an effect on harm, in line with what Black participants showed, but unlike White participants in Studies 2a and 2c. Perhaps outgroup members are more sensitive to written cues than visual cues (Strom et al., 2012). Future research is needed to better understand when and why outgroup members perceive harm when evaluating hate crimes. Notably, the results of Study 1-2b could have been influenced by the fact that low prototypic targets were categorized as Black by the CFD raters more frequently than high prototypic targets. Specifically, this may have resulted in White participants low prototypic condition as biracial, or simply being confused about their race altogether. Further, it is possible that the discrepancies observed between Black and White participants was influenced by Black participants correctly perceiving ingroup members more readily compared than White participants evaluating outgroup members. We addressed this limitation by explicitly labelling all targets as Black in Study 2c. Critically, if a self-identified Black individual is relatively low in Black prototypicality, and is subsequently not categorized as Black by others, it serves to underscore the main problem examined in this work: Such individuals may not be believed when they report hate crimes, may not be seen as harmed by them, and may be excluded from movements, such as #BlackLivesMatter, that are designed to support them.

8.1. Implications

The current research extends the racial prototypicality literature in an applied context and further supports prior research assessing racial

phenotypicality as it relates to social justice movements (e.g., Pejic & Deska, 2024). The current work contributes to the racial prototypicality literature by assessing Black racial identity in the context of the #BlackLivesMatter movement. Further, our studies aimed to examine how racial identity influences judgments of race-based hate crime victimization outcomes. Much of the prior research indicates that highly prototypic Black individuals typically experience significant detriments due to their darker skin colour, such as the heightened activation of negative traits (Maddox & Gray, 2002), denial/disregard of pain and subsequent support (Deska et al., 2020; Trawalter et al., 2012), and harsher criminal sentencing (Blair et al., 2004; Eberhardt et al., 2006). Interestingly, our studies present a unique instance whereby high prototypic Black individuals may actually experience benefits in comparison to their low prototypic Black counterparts. Specifically, our research shows how targets low in Black racial phenotypicality (e.g., lighter skin tone) experience exclusion and negative ramifications compared to their high phenotypic counterparts when focusing on inclusion in the #BlackLivesMatter movement. Although these results are in line with our hypotheses rooted in the sexual assault and prototypicality literature (Goh et al., 2021), they represent a unique pattern when looking at the Black prototypicality literature. Furthering the unique contribution of our research to the literature, our results align with prior research identifying and demonstrating the need to assess racial prototypicality through both the intra- and inter-group lens (Maddox, 2004).

Additionally, the current research provides further support for a greater sensitivity to skin tone bias among ingroups (Maddox, 2004). Black participants in Study 2b and 2c showed heightened sensitivity to skin tone bias and its implications on perceptions of harm following hate crime victimization, but White participants in Study 2a and 2c did not, although, in the verbal vignette design in employed in Study 3, White participants did evaluate harm differently across high and low prototypical targets. Although speculative, perhaps ingroup members are more sensitive to visual cues whereas outgroup members are more sensitive to written cues. Importantly, both ingroup and outgroup members may evaluate both sources of cues. Indeed, ingroup members are not the only individuals involved with social justice organizations such as #BlackLivesMatter. Outgroup members who are allies and/or who are in positions of power may have a stake in hate crime reports and may even directly evaluate these reports. More broadly, although ingroup members often act as gatekeepers for group membership, outgroup members, especially those in socially dominant positions, have had a long history of socially consequential judgments of others based on group membership vis-à-vis perceived prototypicality (e.g., Blair et al., 2002; Eberhardt et al., 2006). Together, beyond replication of this ingroup skin tone bias in the present work, our findings underscore the need to examine evaluations from both intra- and inter-group members.

Finally, our results suggest that low Black phenotypic hate crime victims may be perceived as less credible, and that their reporting (even through a race-specific outlet such as #BlackLivesMatter), is deemed less appropriate compared to high Black phenotypic victims. Accordingly, low Black phenotypic victims may be less likely to report their incidents of race-based hate crime victimization due to skepticism toward their victimhood, as well as perceived exclusion from race-specific support outlets. Conversely, if low Black phenotypic victims do report their hate crime incident, they may feel the need to present their case differently than high phenotypic victims (e.g., be more detailed, make direct connections to their race and the hate crime incident). That is, low Black phenotypic hate crime victims may problematically require greater explanation to prove that they were indeed a victim of a race-based hate crime in order to receive others' belief and support. It is imperative to ensure that race-based hate crime victims receive the adequate support they need, and our results demonstrate that being a non-prototypical victim may pose a barrier in such instances.

8.2. Limitations and future directions

The current work's limitations serve as valuable avenues for future research. First, as we used the CFD's (Ma et al., 2015) norming data to select stimuli, we are unable to confidently disentangle the specific aspects of the stimuli (e.g., skin tone vs. other indicators of phenotypicality) that influenced Black racial prototypicality ratings in our image-based studies. Previous research suggests that judgments of target's racial category membership are influenced by physical cues beyond skin tone such as the broadness of an individuals' lips and/or nose as well as hair texture (Blair et al., 2002). However, we manipulated Black racial prototypicality in Study 3 without the use of images (i.e., text-based descriptions). As hate crime reports are not always associated with images, this additional manipulation also served as a more ecologically valid method given the intended implications of the current study. However, target skin tone variation was the phenotypicality cue in our vignette manipulation. As such, future research should employ additional indicators of Black racial phenotypicality outside of skin tone and image-only based studies.

Further, we did not provide participants with an explicit definition or information regarding what representation meant in relation to the targets and the #BlackLivesMatter movement. As participants were asked to rate how representative of the #BlackLivesMatter movement targets were, and no definition of representation was provided, participants may have understood and applied this definition in various ways. Representation is a multi-faceted construct that can include many different ideas. As such, future research should examine what it means to be represented by a social justice movement, and apply more theoretical control when including this construct.

To help address the issue of the masculine default (Cheryan & Markus, 2020), we used both male and female targets in our studies. However, our research was not designed to examine differences in phenotypicality level as a function of target sex. There is good evidence suggesting that perceptions of race and gender covary (Freeman & Ambady, 2011; Johnson et al., 2012). Thus, future research should expand on the current research program by aiming to examine the complex relationship between race, sex, and other potentially covarying cues (e.g., body size; Wilson et al., 2017) as they relate to judgments of hate crime reporting metrics.

Additionally, participants in all of our studies were exposed to all study conditions. Thus, participants were required to compare each of the conditions they were presented with. As a result, such comparisons may have influenced the results. Although this is a limitation of the research design, targets were presented in a randomized order in all of the studies to minimize concerns about order effects. However, future replication attempts should consider employing between-subjects designs to test whether these results generalize across experimental design.

The current work focuses exclusively on the #BlackLivesMatter movement within the United States. Although present in other parts of the world, it is predominantly a US-based social justice movement. More work is needed to examine how prototypicality impacts who is represented by social justice movements seeking justice for various groups around the world. For instance, colorism is a significant issue in India (Bajwa et al., 2023; Mishra, 2015), where grassroots movements such as Dark is Beautiful seek to raise awareness about colour bias. Japan is recently facing calls for increased recognition of biracial individuals in response to criticism directed at multiracial celebrities such as Naomi Osaka and Ariana Miyamoto. And, as highlighted by movements such as #PapuanLivesMatter in Indonesia, Indigenous peoples around the world face substantially worse outcomes than settlers. But, it remains unclear how variability in category prototypicality across these diverse contexts might influence beliefs about who is represented by those movements, and who is victimized by hate crimes relevant to those groups.

Finally, the current work still focuses narrowly on Black and White targets and perceivers. Although a greater understanding of both ingroup and outgroup perceptions among Black and White individuals is

important, a significant proportion of the existing literature focuses on these populations. Future research needs to examine the phenotypicality cues associated with other racial and ethnic backgrounds. In particular, more work is needed to understand how multiracial people are represented by social justice movements. Multiracial individuals are the largest growing population in the world (Rico et al., 2023), but little work examines phenotypicality within the context of multiracial people (cf. Chen & Hamilton, 2012), let alone how they may or may not be supported when victimized because of their identities.

Open practices

Materials and data for all studies can be found at: https://osf.io/fwq9z/?view_only=fa92dbea1ea7401bb69170bbc4654453.

CRediT authorship contribution statement

Maire L. O'Hagan: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Samantha R. Pejic:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jason C. Deska:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2024.104696.

References

Adams, E. A., Kurtz-Costes, B. E., & Hoffman, A. J. (2016). Skin tone bias among African Americans: Antecedents and consequences across the life span. *Developmental Review, 40*, 93–116. https://doi.org/10.1016/j.dr.2016.03.002

Bajwa, M. J., von Maur, I., & Stephan, A. (2023). Colorism in the Indian subcontinent—Insights through situated affectivity. *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-023-09901-6

Black Lives Matter. (2024, January). About black lives matter. *https://blacklivesmatter.com/about/*.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*(10), 674–679. https://doi.org/10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Judd, C. M., Sadler, M., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83*(1), 5–25. https://doi.org/10.1037/0022-3514.83.1.5

Bryant-Davis, T., Adams, T., Alejandre, A., & Gray, A. A. (2017). The trauma lens of police violence against racial and ethnic minorities. *Journal of Social Issues, 73*(4), 852–871. https://doi.org/10.1111/josi.12251

Chen, J. M., & Hamilton, D. L. (2012). Natural ambiguities: Racial categorization of multiracial individuals. *Journal of Experimental Social Psychology, 48*(1), 152–164. https://doi.org/10.1016/j.jesp.2011.10.005

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review, 127*(6), 1022–1052. https://doi.org/10.1037/rev0000209

Deska, J. C., Kunstman, J. W., Bernstein, M. J., Ogungbadero, T., & Hugenberg, K. (2020). Black racial phenotypicality shapes social pain and support judgments. *Journal of Experimental Social Psychology, 90*, Article 103998. https://doi.org/10.1016/j.jesp.2020.103998

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*(5), 383–386. https://doi.org/10.1111/j.1467-9280.2006.01716.x

Fahs, B., & Swank, E. (2022). Friends or foes? U.S. women's perceptions of racial justice and the black lives matter protests during the COVID-19 pandemic. *Journal of Women, Politics & Policy, 43*(4), 446–462. https://doi.org/10.1080/1554477X.2022.2034429

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*(2), 247–279.

Goh, J. X., Bandt-Law, B., Cheek, N. N., Sinclair, S., & Kaiser, C. R. (2021). Narrow prototypes and neglected victims: Understanding perceptions of sexual harassment. *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes, 122*(5), 873–893. https://doi.org/10.1037/pspi0000260

Goh, J. X., & McCue, J. (2021). Perceived prototypicality of Asian subgroups in the United States and the United Kingdom. *Journal of Experimental Social Psychology, 97*. https://doi.org/10.1016/j.jesp.2021.104201. Article 104201.

Hordge-Freeman, E., & Loblack, A. (2021). "Cops only see the Brown skin, they could care less where it originated": Afro-Latinx perceptions of the #BlackLivesMatter movement. *Sociological Perspectives, 64*(4), 518–535. https://doi.org/10.1177/0731121420961135

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102*(1), 116–131.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Kahn, K. B., Money, E. E. L., & Lake, J. (2024). Whose pain matters? Racial differences in perceptions of emotional pain after fatal police shootings. *Race and Social Problems, 16*, 342–356. https://doi.org/10.1007/s12552-024-09413-1

Keith, V. M., & Herring, C. (1991). Skin tone and stratification in the black community. *American Journal of Sociology, 97*(3), 760–778. https://doi.org/10.1086/229819

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82*(1), 5–18. https://doi.org/10.1037/0022-3514.82.1.5

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review, 8*(4), 383–401. https://doi.org/10.1207/s15327957pspr0804_4

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. https://doi.org/10.1177/0146167202282010

Maddox, K. B., Perry, J. M., & Pagan, J. (2022). Cues and categories: Revisiting paths to racial phenotypicality bias. *Social and Personality Psychology Compass, 16*(8), Article 12699. https://doi.org/10.1111/spc3.12699

Mishra, N. (2015). India and colorism: The finer nuances. *Washington University Global Studies Law Review, 14*(4), 725–750.

Monk, E. P., Jr. (2015). The cost of color: Skin color, discrimination, and health among African-Americans. *American Journal of Sociology, 121*(2), 396–444. https://www.jstor.org/stable/10.1086/682162?origin=JSTOR-pdf.

Monk, E. P., Jr. (2021). The unceasing significance of colorism: Skin tone stratification in the United States. *Daedalus, 150*(2), 76–90. https://doi.org/10.1162/daed_a_01847

Pejic, S. R., & Deska, J. C. (2024). *Narrow prototypes of Asian subgroups in the United States: Implications for the stop Asian hate movement* (Manuscript under review).

Pew Research Centre. (2023). Facts about the U.S. Black population. https://www.pewresearch.org/social-trends/fact-sheet/facts-about-the-us-black-population.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Rico, B., Jacobs, P., & Coritz, A. (2023, June 01). *2020 census shows increase in multiracial population in all age categories*. United States Census Bureau. https://www.census.gov/library/stories/2023/06/nearly-a-third-reporting-two-or-more-races-under-18-in-2020.html#:~:text=The%20Multiracial%20population%20increased%20for,9%2C535%2C241)%20growing%20by%20300%25.

Rosch, E. (1978). Principles of categorization. In A. M. Collins, & E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 312–322). Morgan Kaufmann. https://doi.org/10.1016/B978-1-4832-1446-7.50028-5 (Reprinted from Rosch, E., "Principles of Categorization," in E. Rosch & B. B. Lloyd (Eds.), "Cognition and Categorization," 27-48, Erlbaum, 1978).

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

Strom, M. A., Zebrowitz, L. A., Zhang, S., Bronstad, P. M., & Lee, H. K. (2012). Skin and bones: The contribution of skin tone and facial structure to racial prototypicality ratings. *PLoSONE, 7*(7), Article e41193. https://doi.org/10.1371/journal.pone.0041193

Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial bias in perceptions of others' pain. *PLoS One, 7*(11). https://doi.org/10.1371/journal.pone.0048546. Article e48546.

U.S. Department of Justice. (2023, July). Hate crimes. https://www.justice.gov/hatecrimes/learn-about-hate-crimes.

U.S. Department of Justice. (2023b). *2022 FBI hate crime statistics*. Community Relations Service U.S. Department of Justice. https://www.justice.gov/crs/highlights/2022-hate-crime-statistics.

Vogel, T., Ingendahl, M., & Winkielman, P. (2020). The architecture of prototype preferences: Typicality, fluency, and valence. *Journal of Experimental Psychology: General, 150*(1), 187–194. https://doi.org/10.1037/xge0000798

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of Personality and Social Psychology, 113*(1), 59–80. https://doi.org/10.1037/pspi0000092