



Co-Editors: Benjamin Balas &amp; Jonathan Flombaum

ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/pvis20](http://www.tandfonline.com/journals/pvis20)

# On the influence of implicit race attitudes on explicit trustworthiness judgments: An investigation of the perceivers and targets' race and gender intersection

Alessia Valmori, Erdem O. Meral, Miriam-Linnea Hale, Patrice Rusconi & Marco Brambilla

**To cite this article:** Alessia Valmori, Erdem O. Meral, Miriam-Linnea Hale, Patrice Rusconi & Marco Brambilla (07 Mar 2024): On the influence of implicit race attitudes on explicit trustworthiness judgments: An investigation of the perceivers and targets' race and gender intersection, Visual Cognition, DOI: [10.1080/13506285.2024.2315800](https://doi.org/10.1080/13506285.2024.2315800)

**To link to this article:** <https://doi.org/10.1080/13506285.2024.2315800>



View supplementary material [↗](#)



Published online: 07 Mar 2024.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



# On the influence of implicit race attitudes on explicit trustworthiness judgments: An investigation of the perceivers and targets' race and gender intersection

Alessia Valmori <sup>a</sup>, Erdem O. Meral <sup>b,c</sup>, Miriam-Linnea Hale <sup>d</sup>, Patrice Rusconi <sup>e</sup> and Marco Brambilla <sup>f</sup>

<sup>a</sup>Department of Developmental and Socialization Psychology, University of Padova, Padova, Italy; <sup>b</sup>Department of Social Psychology, Tilburg University, Tilburg, The Netherlands; <sup>c</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands;

<sup>d</sup>Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg; <sup>e</sup>Department of Cognitive Sciences, Psychology, Education and Cultural Studies (COSPECS), University of Messina, Messina, Italy; <sup>f</sup>Department of Psychology, University of Milano-Bicocca, Milan, Italy

## ABSTRACT

People judge strangers' trustworthiness based on their facial appearance, but these judgments are biased. Biases towards Black individuals may stem from implicit pro-White attitudes. However, previous studies have not explored if these effects extend to different members within the same social group, like women instead of men, nor considered the role of the perceiver's multiple social group memberships. Therefore, we investigated how the perceiver's and target's gender influence implicit and explicit biases in trustworthiness judgments of Black and White individuals. Our study included 309 participants, split between Black and White men and women, who completed trustworthiness judgment tasks and the Implicit Association Test (IAT). We found that implicit pro-White attitudes predicted trustworthiness disparities only for White individuals. Interestingly, implicit pro-White attitudes also correlated with explicit pro-men attitudes, but only among male participants. This study underscores the importance of considering multiple social group memberships when studying prejudice in trustworthiness judgments.

## ARTICLE HISTORY

Received 4 October 2022  
Accepted 9 January 2024

## KEYWORDS

Trustworthiness judgments;  
Social biases; Intersectional  
approach; Implicit attitudes;  
Race disparity

People make quick trustworthiness judgments from faces (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Previous work has shown that these trustworthiness judgments can have important consequences such as influencing criminal sentencing decisions (Wilson & Rule, 2015), personnel selection (Gomulya et al., 2017), or the results of political elections (Brizi & Mannetti, 2018), and they are a proxy for one of the fundamental dimensions of face evaluation, that is, valence (Oosterhof & Todorov, 2008). Besides a few exceptions (e.g., Lui et al., 2022; Walker et al., 2011), most research on trustworthiness judgments discusses the universality of such judgments and does so only based on a limited stimuli set (e.g., only faces of White individuals or only male individuals) or a subject pool (e.g., majority White participants). These judgments, however, are not free from bias, especially when judging the faces of individuals with marginalized identities (e.g.,

Charbonneau et al., 2020; Kubota et al., 2013; Stanley et al., 2011).

More recently, there have been numerous calls to incorporate a more intersectional approach in psychological sciences. This would suggest that the effects of multiple group membership create a unique experience that can be inherently different from the additive effects of these group identities (Cole, 2009). Scholars have long been advocating for the adoption of an intersectional approach in the study of multiple identities such as race, gender, sexual orientation or social class (e.g., Hester et al., 2020; Parent et al., 2013; Rosenthal, 2016). To respond to these calls and advance research on facial trustworthiness, in the current study we adopted an intersectional lens. To do so, first we investigated trustworthiness judgments not only towards men but also towards women. Moreover, we recruited an equal number of White and Black individuals (balanced in terms of gender) as

our participants (for a more detailed discussion on the issue, see Else-Quest & Hyde, 2016). We built on previous work investigating how implicit race attitudes influence trustworthiness judgments of Black and White individuals (Stanley et al., 2011). Stanley et al. (2011) reported two studies relying on face photos of Black and White male individuals. The authors tested how the implicit pro-White attitudes measured by the Implicit Association Test (IAT, Greenwald et al., 1998) predicted trustworthiness judgments or economic trust decisions made about the persons in the photos. Their work showed that implicit race attitudes (i.e., an implicit pro-White/anti-Black attitude) predicted race disparity in trustworthiness judgments and trust decisions in an economic game favouring White targets over Black targets.

Our study was spurred by two research questions. First, does the disparity in trustworthiness judgments favouring White individuals also apply to women's faces and not only men's faces? Second, do Black participants exhibit the same bias as White participants with regards to the positive relationship between pro-White attitudes and disparity in trustworthiness judgments?

### ***Implicit race attitudes and trustworthiness judgments***

A good deal of work highlighted how trustworthiness impressions from faces form rapidly (Olivola & Todorov, 2010; Willis & Todorov, 2006) and spontaneously (Engell et al., 2007; Klapper et al., 2016). These quick trustworthiness judgments may not be free from bias. There is evidence suggesting that one's attitudes towards the race of the target person (i.e., the person being evaluated) influences trustworthiness judgments. For instance, one study showed that the level of implicit bias towards an ethnic outgroup positively relates to having a mental representation of that group as less trustworthy (Dotsch et al., 2008). In another study, it was observed that implicit pro-White attitudes moderated people's explicit trustworthiness ratings and trust decisions for Black individuals (Stanley et al., 2011). Results suggested that, to the extent that individuals held implicit pro-White attitudes, they were more likely to favour White individuals over Black individuals in trustworthiness judgments and trusting behaviors. Similarly, results of another study suggested

that pro-White attitudes were predictive of accepting fewer offers from Black (vs. White) targets (also presented via face photos) in an economic game (Kubota et al., 2013). In a similar vein, another study found that White participants who had a high level of pro-White implicit bias perceived anger in the faces of Black targets more readily, but this pattern was not observed for White faces (Hugenberg & Bodenhausen, 2003). Taken together, the evidence shows that trustworthiness judgments of Black individuals may be biased to the extent that one holds pro-White/anti-Black implicit attitudes.

### ***Current study***

Scholars have studied perception and social judgments with an intersectional lens in the past. For instance, prior work tested how race and gender influence categorization (Goff et al., 2008) and perception (Livingston et al., 2012; Opie & Phillips, 2015). This work reveals the relative invisibility of Black women (Sesko & Biernat, 2010). More akin to the efforts in the present study, studies with an intersectional approach also tested whether stereotypes about social groups apply to all identities in a given group. For example, one study revealed that the implicit racial bias associating young Black (vs. White) men with criminality and violence applies to old black men as well (Lundberg et al., 2018). Taken together, a growing body of work highlights the importance of adopting an intersectional approach in understanding social judgments (e.g., Petsko et al., 2022).

The current project builds on this body of work by applying an intersectional lens to the study of how implicit biases influence trustworthiness judgments. While prior has tested how implicit biases impact trustworthiness judgments (Stanley et al., 2011) and how social judgment is shaped by an intersectional lens (Lundberg et al., 2018), the combination of these two research streams seems to have received less empirical attention. Do implicit biases similarly impact trustworthiness judgments about different members of the same social groups (e.g., Black men and Black women)? Do members of different social groups (e.g., White and Black men/women) exhibit similar levels of race disparity in trustworthiness judgments in relation to their implicit biases?

To answer such questions, we aimed at testing the role of gender in shaping disparities in

trustworthiness judgments made about Black and White individuals. Our design was based on Study 1 by Stanley et al. (2011), in which participants rated photos of Black and White individuals on trustworthiness and completed an IAT task. Stanley et al. (2011) did not incorporate female targets in their study likely due to the bias under scrutiny being mostly associated with Black men (this practice is also observed in other research on the topic: Charbonneau et al., 2020; Kubota et al., 2013). To answer questions that are more intersectional in nature, we incorporated an equal number of male and female targets (Black and White). Two competing hypotheses may be raised on how target gender would interact with target race to produce the race disparity in trustworthiness judgments. One possibility could be that racial disparity in trustworthiness judgments would be observed only for male targets (H1). Alternatively, racial disparity in trustworthiness judgments could also be observed for female face photos (H2). This hypothesis would be supported by findings suggesting that Black women are masculinized and people differentiate Black women and Black men to a lesser extent than they do with White men and White women (Coles & Pasek, 2020; Goff et al., 2008). Accordingly, such an intersectional invisibility of Black women would suggest that the same racial disparity would be observed for male and female targets.

Another way we aimed to adopt an intersectional lens was by incorporating the perceiver's race (Black vs. White) and gender (Man vs. Woman) into our design. We have ensured this by recruiting an equal number of Black and White participants and an equal number of men and women as participants. This allowed us to explore whether the race disparity observed in Study 1 by Stanley et al. (2011) would be moderated by the perceiver's race and gender (RQ1). Stanley et al. (2011) suggested that the perceiver's race did not have a significant influence on race disparity. This finding is in line with past research suggesting that both Black and White individuals exhibit similar levels of implicit race bias (Kahn & Davies, 2011). However, previous work has also suggested that individuals have a more positive attitude towards faces similar to theirs and, in turn, evaluate these faces as more trustworthy (Platek et al., 2009) suggesting that perceiver's race may influence the race disparity effect (such that individuals only display the effect for other-race faces). Thus, we

propose that investigating trustworthiness judgments with an intersectional lens could contribute to better understanding these seemingly conflicting results. In a similar vein, we also explored whether implicit pro-White attitudes would relate to a gender disparity in trustworthiness judgments (RQ2).

## Method

### Participants and design

Participants were recruited online through Prolific (a crowdsourcing platform; each participant received € 2,00 as compensation for their efforts). Previous work has shown that data from 26 independent raters attain stable trustworthiness ratings (Hehman et al., 2018). We therefore also wanted to recruit enough participants so that each of the 164 photos we used was rated by 26 participants. We chose to enrol 316 participants, aiming for an approximate distribution of 100 raters for each image set. This decision was made to secure a well-balanced representation of gender and race within the sample, targeting approximately 25 raters per social group. Nevertheless, as shown in Table 1, only 309 participants were included in the final analysis.

One participant's data were removed from the analyses because they did not complete the survey, six participants' data were removed because they did not identify as either Black or White and three participants' data were removed as they did not identify as either man or woman. Participants were balanced for their gender ( $N_{\text{Men}} = 158$ ,  $N_{\text{Women}} = 151$ ) and race ( $N_{\text{Black}} = 149$ ,  $N_{\text{White}} = 160$ ), their mean age was 25.59 ( $SD = 7.60$ ). Considering the final analyzed sample, each photo was rated by about 100 participants.

The study used a 2 (target's race: Black vs. White)  $\times$  2 (target's gender: men vs. women) within-subject design.

### Procedure and materials

The entire study was anonymously administered online via Qualtrics. We took all the neutral photos

**Table 1.** Distribution of the 309 final participants across the three sets of images for the explicit judgment task.

	Set 1		Set 2		Set 3	
	Black	White	Black	White	Black	White
Men	22	32	28	22	26	28
Women	28	22	16	30	29	26

from the Chicago Face Database (CFD, Ma et al., 2015). In total, we used 196 photos, of which 88 were photos of Black, 88 were photos of White, 10 were photos of Asian individuals and 10 were photos of Latinx individuals. Each group of images was balanced for the gender of the target. One-hundred-sixty-four images (i.e., 72 Black targets, 72 White targets, 10 Asian targets and 10 Latinx targets) were employed in the explicit judgment task and the remaining 32 images (i.e., 16 Black targets and 16 White targets) were shown in the IAT. Importantly, we analyzed the normative data for Black and White targets of the selected images of the explicit task and the IAT retrieved from Ma et al. (2015).<sup>1</sup> We performed two different ANOVAs considering the gender and the ethnicity of the targets as between-subject factors. When the images selected for the explicit judgments were considered, it emerged that CFD raters considered these targets differing for the Trustworthiness dimension as a function of their gender,  $F(1,140) = 10.02$ ,  $p = .002$ ,  $\eta^2_p = .07$ . Generally, women targets were considered more trustworthy compared to men. Moreover, the interaction between the Gender and the Ethnicity of the targets emerged,  $F(1,140) = 11.64$ ,  $p < .001$ ,  $\eta^2_p = .08$ . Although no difference emerged for Black targets, when White targets are considered, men are judged as less trustworthy compared to women. Table 2 reports the normative data of the selected images for the explicit judgment task for the Trustworthiness dimension.

When the normative data of the images selected for the IAT were considered, only the main effect of the Ethnicity emerged,  $F(1,28) = 6.24$ ,  $p = .02$ ,  $\eta^2_p = .18$ . Overall, White targets were considered less trustworthy than Black targets. Table 3 reports the normative data of the selected images for the IAT for the Trustworthiness dimension. The *a priori* imbalance of our database could account for participants' explicit and implicit ratings about the trustworthiness of the targets. However, our main aim was to

investigate participants' attitude in relation to their demographics. For this reason, we did not consider that the imbalance would prevent us from testing our hypotheses. More information about the normative data of the selected targets is reported in the Supplementary materials (S1 and S2).

First, participants completed an explicit ratings task. In order not to induce unnecessary fatigue on our participants, we created smaller subsets that included an equal number of Black, White, male, and female faces. Each subset also incorporated photos of Asian and Latinx individuals in order not to make the comparison between Black and White faces salient. In the end, we had three subsets with 68 photos each with photos of 24 Black, 24 White, 10 Latinx, and 10 Asian individuals. More data about the results for these stimuli are reported in the Supplementary materials. Since our focus was on the Black and White individuals, we used the same photos of Latinx and Asian individuals across the three subsets. Participants were then randomly assigned to rate only one of the three subsets.

Participants were told that they were involved in a study about filters in social media and that they would evaluate faces either with filters or without filters. In reality, all participants saw the photos in their original format (i.e., the no-filter condition in the cover story). Participants first rated a trial set of 10 photos different from the ones of the experiment set to become familiar with the procedure. Afterwards, they moved on to the actual set and rated all the photos in one of the three subsets. They rated each face (68 in total) on trustworthiness (1 = *not trustworthy at all*, 9 = *very trustworthy*). Differently from Stanley et al. (2011), the photo remained visible until participants expressed their judgment (Figure 1).

Next, participants completed a Black and White race IAT implemented online with the iatgen tool for Qualtrics surveys (Carpenter et al., 2019), in

**Table 2.** Descriptive statistics and pairwise comparisons for the trustworthiness dimension of the selected images for the explicit judgment task.

	Black M (SD)	White M (SD)	Total M (SD)
Men	3.47(.31) <sub>a</sub>	3.25 (.33) <sub>b</sub>	3.36 (.34) <sub>d</sub>
Women	3.46 (.35) <sub>a</sub>	3.60 (.28) <sub>c</sub>	3.53 (.34) <sub>e</sub>
Total	3.47 (.33) <sub>f</sub>	3.42 (.35) <sub>f</sub>	

Note: Different letters indicate significant differences across rows and columns. Bonferroni post-hoc comparison for the interaction, all  $ps < .01$ .

**Table 3.** Descriptive statistics and pairwise comparisons for the trustworthiness dimension of the selected images for the IAT.

	Black M (SD)	White M (SD)	Total M (SD)
Men	3.62 (.50) <sub>a</sub>	3.35 (.38) <sub>a</sub>	3.49 (.45) <sub>b</sub>
Women	3.81 (.45) <sub>a</sub>	3.35 (.30) <sub>a</sub>	3.58 (.44) <sub>b</sub>
Total	3.72 (.47) <sub>c</sub>	3.35 (.33) <sub>d</sub>	

Note: Different letters indicate significant differences across rows and columns.

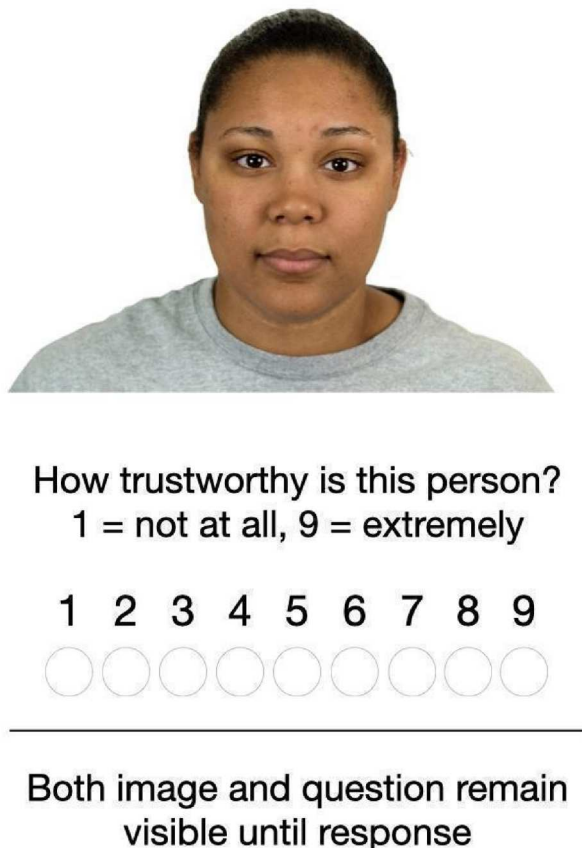


which we measured the strength of the association between the target (Black or White) and the pleasant (vs. unpleasant) domain. Participants completed the IAT twice, once with male pictures and once with female pictures, counterbalancing the presentation order. For each task, participants completed seven categorization blocks by pressing the instructed buttons on their keyboard. In each trial, either an image or a word appeared. Images displayed Black vs. White targets. Words referred to pleasant (i.e., "wonderful," "friend," "laughter," "love," "peace," "pleasure," "joy," and "happy") or unpleasantness concepts (i.e., "nasty," "terrible," "war," "awful," "failure," "agony," "evil," and "horrible"). Participants had to categorize the stimulus into the correct category (i.e., "Black" vs. "White" or/and "unpleasantness" vs. "pleasantness"; labels were always shown in the upper corners of the screen) by pressing the correct key button (i.e., "E" vs. "I"). During the critical trials, stimuli alternated between images of Black vs. White targets or words of pleasantness vs. unpleasantness. The intertrial interval was 250 ms. When participants made errors, they had to correct their

answers before proceeding. The entire task lasted approximately 5 min. Specifically, Block 1 was a practice block (20 trials) of only images of Black vs. White targets. Participants had to correctly sort the White targets images with the button "E" and the Black targets images with button "I." Block 2 was another practice block (20 trials), in which only words of pleasantness vs. unpleasantness were presented. Participants had to correctly categorize the unpleasant words with the button "E" and the pleasant words with button "I." Next, two combined blocks were presented that were crucial for examining participants' implicit White favouritism. In Block 3, 20 practice trials were administered using both images and words. Participants had to categorize White targets images and pleasant words by pressing the "E" button in the keyboard. Similarly, they had to categorize Black targets images and unpleasant words by pressing the key button "I." Block 4 was identical to Block 3 and participants completed 40 critical trials (scoring uses data from Block 3 and Block 4). Thereafter, another practice block was presented (Block 5), consisting of the 40 trials of images of Black vs. White targets with the sides reversed (i.e., "E" key-button for Black targets and "I" key-button for White targets). This helped counterbalance left-right associations learned in the early blocks. Finally, participants repeated the combined block with the categories in their reversed positions. As before, this is divided into 20 practice trials (Block 6) and 40 critical trials (Block 7). Following the original version of the IAT (Greenwald et al., 1998), other three counterbalanced versions of the task were created.

Data in the combined blocks (Block 3 + Block 4 and Block 6 + Block 7) were then analyzed following Greenwald et al.'s (1998) algorithm. A standardized difference score (D score) was calculated for each participant, indicating in which condition (i.e., White favouritism vs. Black favouritism) participants had a better performance. A D score of 0 indicates no difference in attitude towards Black vs. White targets; a positive score indicates that one had a more positive attitude towards White targets rather than Black targets; and a negative score indicates that one had a more positive attitude towards Black targets rather than White targets.

At the end of the study, participants were thanked and debriefed by describing the real aims of the experiment.



**Figure 1.** Diagram of a single trustworthiness trial.

## Results

### Implicit judgments

First, we looked at participants' implicit race attitude by analyzing the IAT associations both for men and women pictures. Yet, due to a technical error in the women targets' IAT, we were only able to use the data from 77 participants<sup>2</sup> ( $N_{\text{Men}} = 36$ ,  $N_{\text{Black}} = 40$ ,  $M_{\text{age}} = 25.58$ ,  $SD_{\text{age}} = 7.47$ ). However, as the implicit judgments for men and women targets correlated,  $r(75) = .56$ ,  $p < .001$ , we decided to compute a unique score for the participants who completed both the men and women targets' IATs. The final IAT  $D$  score was obtained by following Greenwald, Nosek, & Banaji (2003). The mean of the  $D$  score was .28 ( $SD = .43$ ) indicating that participants generally showed a pro-White attitude compared to no bias,  $t(301) = 11.53$ ,  $p < .001$ ,  $d = .66$ . More importantly, we performed a two-way ANOVA on the IAT  $D$  score including the gender and the race of the participants as variables varying between-subjects. The main effect of participants' race emerged,  $F(1,298) = 15.33$ ,  $p < .001$ ,  $\eta^2_p = .28$ . Pro-White bias was observed among White participants ( $M = .50$ ,  $SD = .37$ ), but not among Black participants ( $M = .05$ ,  $SD = .37$ ). Importantly, Black participants did not show a pro-Black bias either. Similarly, a main effect of gender emerged,  $F(1,298) = 6.77$ ,  $p = .01$ ,  $\eta^2_p = .02$ . The positive bias towards White targets was more pronounced for men ( $M = .34$ ,  $SD = .44$ ) than women participants ( $M = .23$ ,  $SD = .41$ ).

### Explicit judgments

Second, we analyzed participants' explicit trustworthiness ratings. Faces were rated on average 5.21 (range: 1–9,  $SD = 2.02$ ,  $N = 144$ ), indicating that generally participants considered the faces significantly more trustworthy than the average midpoint of 5.00,  $t(14,831) = 12.43$ ,  $p < .001$ ,  $d = .10$ . We estimated different linear mixed models employing the package *lme4* in RStudio V1.3.1093 (Bates et al., 2015). We started estimating a minimal model which included only the random intercepts of participants and pictures. We then added each variable in the subsequent models until finding a model that significantly improved the goodness of fit compared to the previous ones and best described the data compared to more complex

models. By adopting this approach, we were able to investigate the support for H1, H2, and RQ1. The model that best predicted our data included the pro-White implicit judgment (IAT), race and gender of the target, their interactions, and the interaction between the gender of the pictures and the gender of the participants, as well as the interaction between the race of the pictures and the race of the participants as fixed factors. Participants and targets IDs were still included as random factors modelled as intercepts. The R formula that describes this model is:

```
trustworthinessRating ~ IAT + targetRace
+ targetGender + targetRace:targetGender
+ participantGender:targetGender
+ participantRace:targetRace
+ (1|participantID) + (1|targetID).
```

The Marginal  $R^2$  of the model was .03 and the Conditional  $R^2$  of the model was .33. The ICC for the random factors was .31. Results showed that the main effect of race of the person portrayed in the picture ( $B = -.81$ ,  $SE = .14$ ,  $t = -5.78$ ,  $\eta^2_p = .16$ , 95% [CI] =  $[-1.08, -.53]$ ) and the gender of the participant ( $B = -.28$ ,  $SE = .12$ ,  $t = -2.40$ ,  $\eta^2_p = .01$ , 95% [CI] =  $[-.51, -.05]$ ) were significant predictors of trustworthiness ratings. More specifically, participants rated Black targets ( $M = 5.40$ ,  $SD = 1.97$ ) as more trustworthy than White targets ( $M = 5.01$ ,  $SD = 2.06$ ), despite there was not an a priori significant difference between Black and White targets' perceived trustworthiness (see Table 2). This result contradicts previous findings by Stanley et al. (2011), in which no difference between ratings of White and Black targets emerged. In addition, overall male participants ( $M = 5.28$ ,  $SD = 2.04$ ) rated the targets as more trustworthy compared to female participants ( $M = 5.13$ ,  $SD = 2.00$ ). Moreover, we also observed an interaction between gender and race of the target ( $B = .42$ ,  $SE = .20$ ,  $t = 2.17$ ,  $\eta^2_p = .03$ , 95% [CI] =  $[.04, .81]$ ). As shown in Figure 2, although participants rated Black men ( $M = 5.29$ ,  $SD = 1.93$ ) more trustworthy (Bonferroni *post-hoc* comparison  $< .001$ ) than White men ( $M = 4.69$ ,  $SD = 2.07$ ), this difference was no more significant for Black and White women (Black:  $M = 5.51$ ,  $SD = 1.99$ ; White:  $M = 5.34$ ,  $SD = 1.99$ ), Bonferroni *post-hoc* comparison = .54. However, the effect of the interaction between gender and race may be

influenced by the *a priori* imbalance of our database (i.e., greater perceived trustworthiness for Black men compared to White men and for White women compared to Black women) and therefore reproduce the results obtained by Ma et al. (2015).

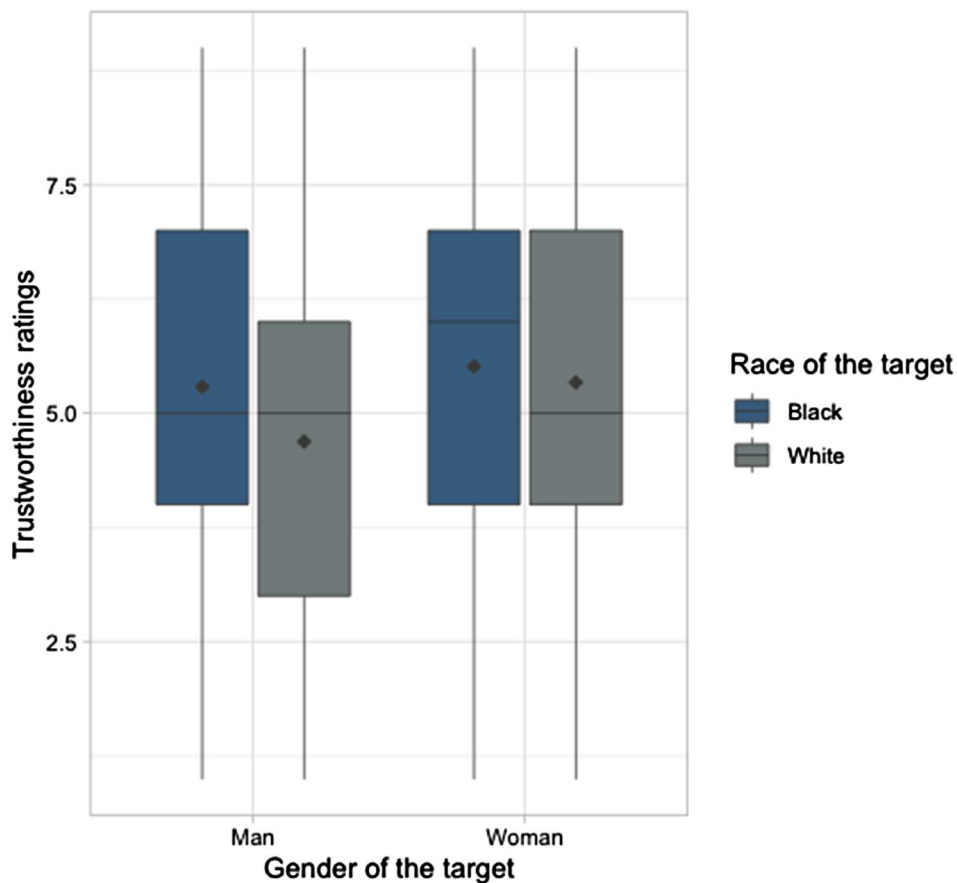
Importantly, also the interaction between the pro-White implicit judgments resulting from the IAT and the race of the target emerged,  $B = .54$ ,  $SE = .10$ ,  $t = 5.44$ ,  $\eta^2_p = .003$ , 95% [CI] = [.35, .74]. A simple slope analysis revealed that the pro-White implicit attitude positively predicted the explicit judgments of trustworthiness for White targets ( $B = .41$ ,  $SE = .16$ , 95% [CI] = [.09, .73]), but this association was no more statistically significant for Black targets ( $B = -.12$ ,  $SE = .16$ , 95% [CI] = [-.43, .20]), *Tukey HSD post-hoc test*  $p < .001$ . Figure 3 shows this interaction.

Similarly, as shown in Figure 4, the interaction between the pro-White implicit judgments resulting from the IAT and the gender of the target emerged,  $B = -.28$ ,  $SE = .09$ ,  $t = -3.00$ ,  $\eta^2_p = .001$ , 95% [CI] = [-.46, -.10]. A simple slope analysis

suggested that there was a tendency for pro-White implicit attitude to be positively predictive of more positive explicit judgments of trustworthiness for men targets ( $B = .29$ ,  $SE = .16$ , 95% [CI] = [-.02, .61]), however no tendency emerged for women ( $B = .001$ ,  $SE = .16$ , 95% [CI] = [-.31, .32]), *Tukey HSD post-hoc test*  $p < .001$ .

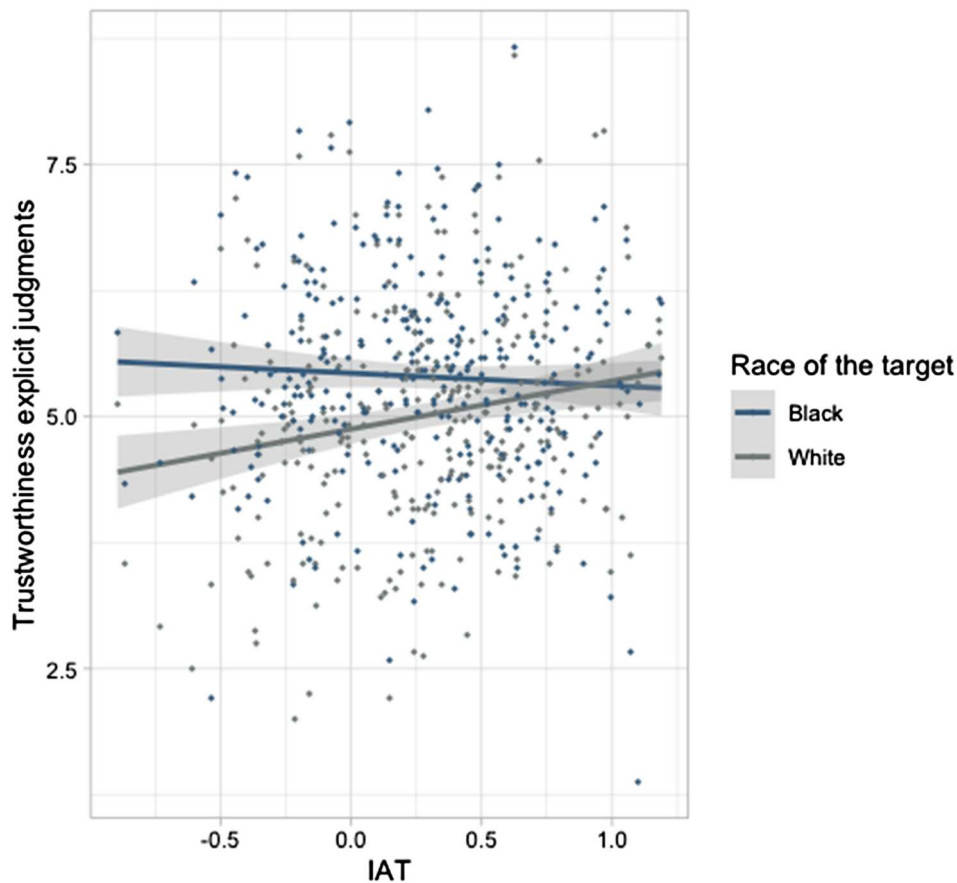
Importantly, these results suggest that generally an implicit pro-White attitude is associated with higher trustworthiness ratings for White and male targets, however this does not reflect in more negative explicit judgments for Black and female targets.

Finally, the two-way interaction between the gender of the target and the gender of the participants emerged ( $B = .28$ ,  $SE = .06$ ,  $t = 5.310$ ,  $\eta^2_p = .002$ , 95% [CI] = [.17, .39]). Figure 5 illustrates that, although due to the initial imbalance of our materials female targets were considered more trustworthy than male targets, this difference was more pronounced for female participants (female targets:  $M = 5.43$ ,  $SD = 1.97$ ; male targets:  $M = 4.83$ ,  $SD = 1.98$ , Bonferroni



**Figure 2.** Interaction between gender and race of the target. The point and the line depicted in each boxplot correspond to the mean and the median, respectively.





**Figure 3.** Interaction between IAT and race of the target. Every data point on the chart represents an individual participant.

*post-hoc* comparison  $< .001$ ) compared to male participants (female targets:  $M = 5.42$ ,  $SD = 2.01$ ; male targets:  $M = 5.14$ ,  $SD = 2.06$ , Bonferroni *post-hoc* comparison = .01).

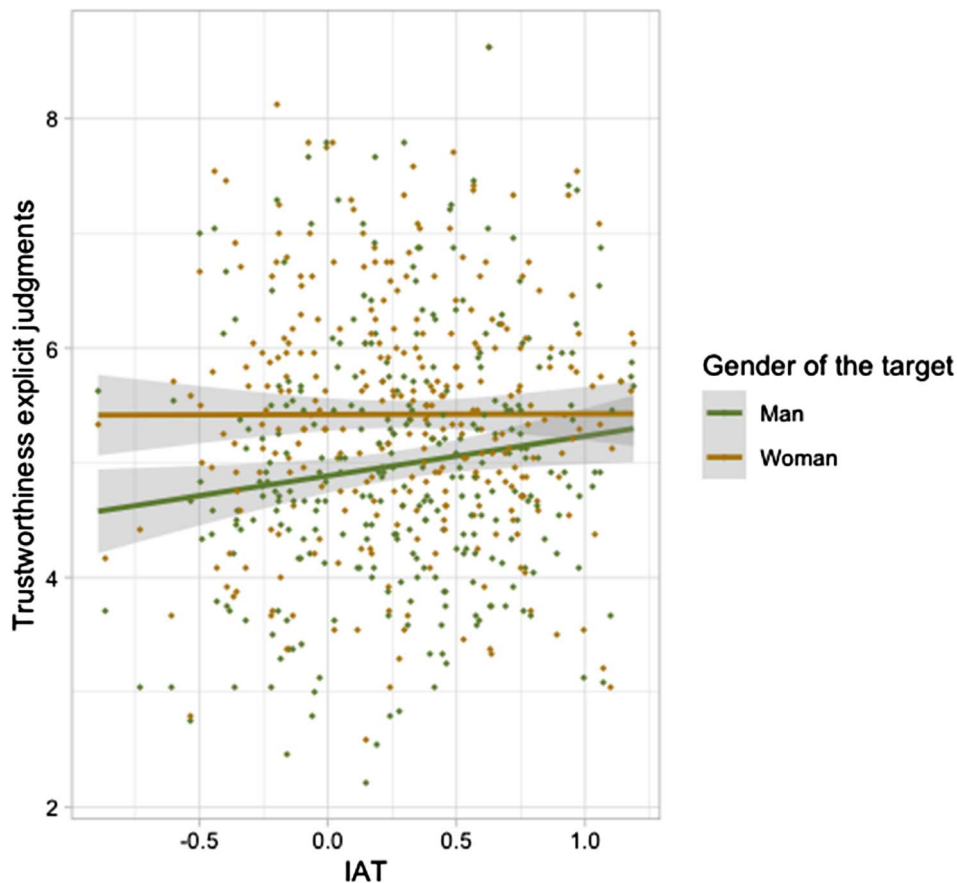
No other significant effect emerged from the analysis,  $ps > .07$ .

### **Correlation between explicit and implicit judgments**

Finally, consistently with Stanley et al.'s (2011) analyses, we correlated the individual differences in IAT score (IAT D score) with individual differences in rating disparity. Following the steps in Stanley et al. (2011), we computed the rating disparity score by subtracting the trustworthiness judgments of Black targets from trustworthiness judgments of White targets. Thus, a positive disparity score indicates that a given participant rated White targets as more trustworthy than Black targets (same direction as the IAT score interpretation). In line with Stanley et al. (2011), the pro-White attitudes correlated

significantly with race disparity in trustworthiness judgments,  $r(300) = .31$ ,  $p < .001$  (Figure 6). We also investigated whether this relationship was observed for both Black and White participants (Figure 7). The positive correlation remained significant only when the White participants were considered,  $r(153) = .31$ ,  $p < .001$ . The correlation was not statistically significant for the Black participants,  $r(145) = .15$ ,  $p = .08$ .

Finally, we extended Stanley et al.'s (2011) analyses by also computing a gender disparity score. It was obtained by subtracting the trustworthiness judgments of women targets from trustworthiness judgments of men targets. Therefore, positive gender disparity scores would suggest higher ratings on the trustworthiness dimension for men targets compared to women. Results shown in Figure 8 revealed that the implicit pro-White attitude positively correlated with an explicit pro-men attitude,  $r(300) = .16$ ,  $p = .005$ . The positive correlation remained significant only when men participants were considered,  $r(151) = .17$ ,  $p = .03$ , but not when only women participants were included in the analysis,  $r(147) = .11$ ,  $p = .20$  (Figure 9).



**Figure 4.** Interaction between IAT and gender of the target. Every data point on the chart represents an individual participant.

## Discussion

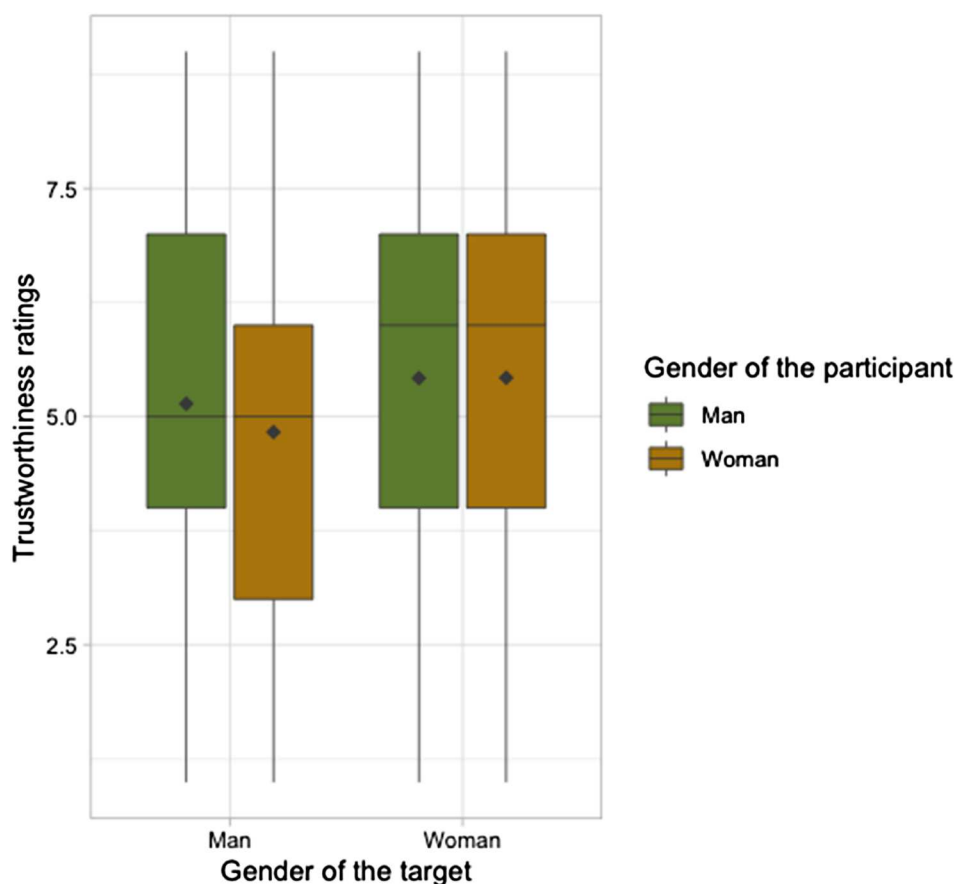
In the current study, we investigated the relationship between implicit bias towards a social group (Black individuals) and explicit judgments of trustworthiness of a member of that social group. We overcame the limitation of previous studies (e.g., Stanley et al., 2011), in which there was a prevalence of White participants in the sample, by incorporating equal numbers of White and Black participants. Filling a gap in the literature, we also diversified the stimulus material by including female faces in addition to male faces.

In our sample, implicit pro-White/anti-Black attitudes predicted higher trustworthiness ratings, but only for White targets, whereas the implicit attitude did not predict the explicit judgments on the trustworthiness dimension for Black targets. Interestingly, participants perceived the photos of Black individuals as more trustworthy than the photos of White individuals in the explicit ratings, despite their a priori trustworthiness was not significantly different (see Table 2).

Additionally, greater pro-White implicit attitude correlated with a higher gender disparity rate, specifically an explicit pro-men attitude for men participants.

We showed that the race disparity effect, as suggested by Stanley et al. (2011) and in line with H1, does not extend to female targets. That is, the implicit pro-White attitude (measured with male face photos in the original study) does not predict a race disparity in trustworthiness judgments of female face photos. With regards to our research question (RQ1), we found that, in contrast to Stanley et al. (2011), only White participants in our sample demonstrated the race disparity in trustworthiness judgments but not Black participants, indicating a moderating effect of participant's race.

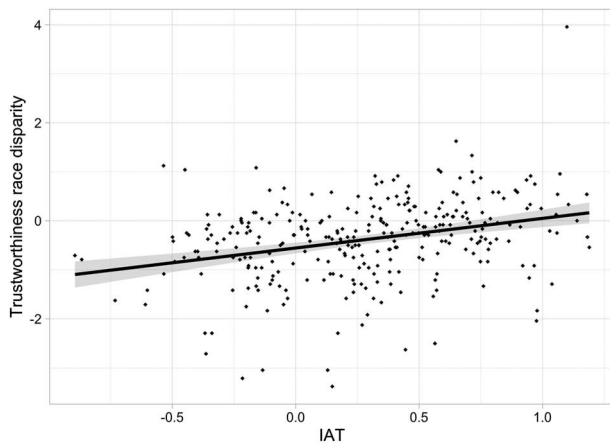
The higher trustworthiness ratings of Black targets are different from the original study which found no difference between Black and White face photos in terms of explicit trustworthiness judgments (Stanley et al., 2011). However, in their discussion the authors did suggest taking implicit bias into account



**Figure 5.** Interaction between gender of the target and gender of the participants. The point and the line depicted in each boxplot correspond to the mean and the median, respectively.

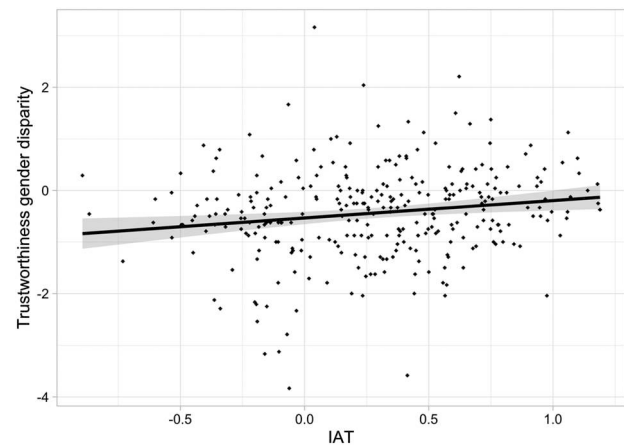
may be key. Another study investigating the relationship between trustworthiness judgments and criminal sentencing by relying on the same database found the same pattern as the current study (Wilson & Rule, 2015), as did the CFD norming data (Ma et al., 2015): Black targets were rated more trustworthy than White targets. However, in our sample this difference was only significant for male faces. The *a priori* norming data showed greater trustworthiness of White female targets than Black male targets, which could explain why the effect was found only for male targets (whose *a priori* norming data showed greater trustworthiness for Black targets than White targets). While this difference can also be found in the norming data of the CFD (see above, Table 2) and it should arguably be considered when interpreting the results, our main aim was to investigate the relationship between the IAT scores and the explicit judgments and not the explicit judgments themselves. The gender–race interaction results also do not entail interactions with any other demographic

variables, so the *a priori* difference in the norming data may account for this result. Thus, we do not see this difference as a major concern. Earlier research indicates that implicit and explicit bias are not correlated if the outgroup is normatively protected (Franco & Maass, 1999). Moreover, studies suggest that implicit and explicit attitudes may show different strengths of correlation as a function of the domain and that this association may be subjected to changes over time (Oswald et al., 2013; Schimmack, 2021). However, Charlesworth and Banaji (2019), suggested that race showed one of the strongest positive correlations between explicit and implicit attitudes, which also revealed to be stable across time. Moreover, a recent study by Stevens and Shriver (2024) suggests that shame may be a reason why people avoid questioning or recognizing their own implicit racial biases which could also partially explain the differences between our findings concerning explicit judgments and implicit racial bias of White participants.



**Figure 6.** Correlation between trustworthiness race disparity judgments and IAT scores. Every data point on the chart represents an individual participant.

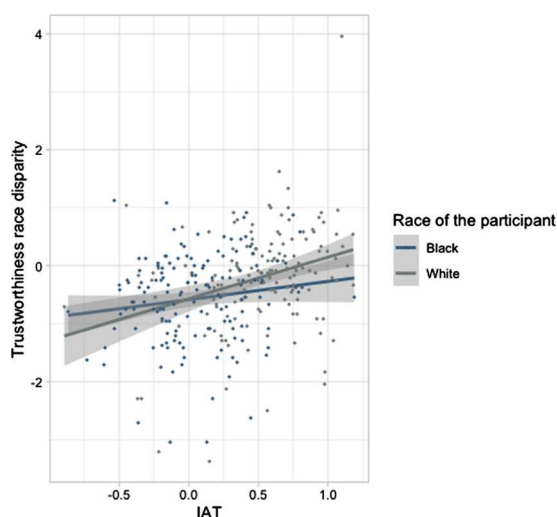
Alternatively, Sawyer and Gampa (2018) found that from 2009 to 2016 during the Black Lives Matter social movements, both implicit and explicit attitudes towards Black persons have changed increasingly and significantly, while biases and stereotypes declined. A similar shift may have occurred in the recent years after the resurgence of the Black Lives Matter movement in 2020 leading to an increase of public, political, scientific and social debates about the social justice movement (Deflem, 2022; Dunivin et al., 2022). That said, we offer this alternative explanation as a speculative one especially given that we did not limit our data collection to the United States. Thus, investigating such a link between



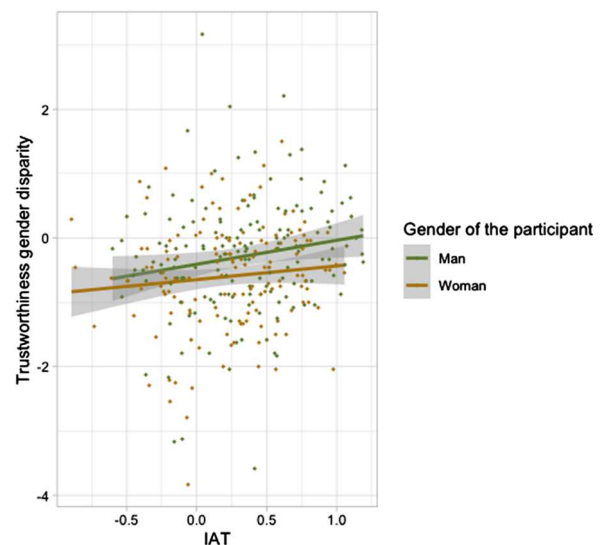
**Figure 8.** Correlation between trustworthiness gender disparity judgments and IAT scores. Every data point on the chart represents an individual participant.

complex sociocultural and psychological phenomena may require future work.

Our extension of the original study responds to a call to incorporate intersectionality in the study of social perception (e.g., Hester et al., 2020; Parent et al., 2013; Rosenthal, 2016). In fact, our findings corroborate the importance of an intersectional lens. Diversifying stimulus material, as well as samples, is not only important for the generalizability of findings. It respects the natural reality of intersecting identities which is crucial in discovering effects unique to any number of intersectional identities. Indeed, the intersectional approach allowed us to



**Figure 7.** Correlation between trustworthiness race disparity judgments and IAT scores separated for Black and White participants. Every data point on the chart represents an individual participant.



**Figure 9.** Correlation between trustworthiness gender disparity judgments and IAT scores separated for men and women participants. Every data point on the chart represents an individual participant.

identify new effects that were hidden in the previous literature. Notably, the main finding that resulted from this methodological rationale is a lack of significant correlation between implicit pro-White attitudes with both the race disparity score for Black participants and the gender disparity score for women participants. While our study shows the importance in expanding research in this way, it is still limited to Black and White and male and female identities respectively. Future research should aim to expand this even further, by incorporating a broader age range, more diverse ethnicities, as well as more diverse gender identities, including non-binary genders. Especially considering the rise of anti-Asian discrimination following the COVID-19 virus from 2020 onward (Lui et al., 2022), it would be interesting to investigate how this impacted trustworthiness judgments as compared to other ethnicities. For instance, Vacchini et al. (2021) reported a higher moral attribution for taxi drivers who rejected foreign clients when their nationality was connected with the geographical locations in which COVID-19 spread the most (i.e., Chinese and Italian) compared to when the nationality was not revealed.

Although we were not able to use the data from all participants for ratings of images depicting women from the IAT, we recommend that future research focus on implicit gender bias in association with race and trustworthiness judgments. Implicit associations between racial groups and positive- or negatively-valenced stimuli have long been the focus of research and the strength of the predictive value of this association is an ongoing debate (e.g., Calvert et al., 2022; Carlsson & Agerström, 2016; Charlesworth et al., 2022; Greenwald et al., 2015; McConnell & Leibold, 2001; Oswald et al., 2013, 2015; Schimmack, 2021). We contribute to this line of work by showing how implicit pro-White attitudes are predictive of trustworthiness judgments for White participants but not for Black participants. We furthermore showed that gender and race can interact when it comes to trustworthiness judgments. Therefore, we recommend future work investigating the relationship between implicit racial attitudes and other judgments or behaviors to incorporate an intersectional approach.

While we have balanced our material, as well as our sample in this study with regards to the number of Black and White and male and female participants,

we did not inquire about the quality and salience of our participants' racial identity, which may also have a significant impact on ingroup vs. outgroup judgments (Charnysh et al., 2015; McCoy & Major, 2003). This would be interesting to incorporate into future research. Furthermore, while we did consider Black vs. White images in our material, we did not account for the different levels of prototypicality, for example colorism, the discrimination based on skin tone, favouring lighter complexions. How dark or light the skin color of the images depicting Black persons is in the stimulus material may have a significant impact on how they are perceived and judged. The intersectionality of sexism and colorism has been especially researched and shown to be impactful for several social outcomes for Black women and men (Barideaux et al., 2021; Mathews & Johnson, 2015; strmic-pawl et al., 2021). Research by Livingston and Pearce (2009) indicates that additionally, factors such as physical structure and perceived warmth of Black faces may diminish the perceived stereotypical threat of these faces. These factors should be considered in future research focusing on judgments and perceptions of images showing Black persons.

Previous research has shown that intersectional effects of race and gender biases often have practical implications in situations where judgments of trust and trustworthiness are important, such as court judgments (Collins & Moyer, 2008; Steffensmeier et al., 2017) or factors such as hiring, leadership perceptions, work experiences, sexual harassment and legal consequences for women in the workplace (Rosette et al., 2018). This is in line with our findings concerning the distinct effects of intersectionality and further speaks to the importance of incorporating it in research, as it has relevant consequences for everyday life.

## Conclusion

In the current study, we investigated how race, gender, and perceiver's implicit attitudes towards race relate to explicit trustworthiness judgments of Black and White targets. Our findings revealed that White perceivers exhibited an implicit pro-White attitude and Black participants did not. Additionally, this implicit pro-White attitude was positively related with higher trustworthiness ratings for White targets and was not significantly related with ratings of Black targets. Importantly, in contrast to previous work



(Stanley et al., 2011), our results suggested that the implicit pro-White attitudes only predicted racial disparity in trustworthiness judgments for White, but not Black, participants and may not apply to female targets. Taken together, our results show the complexity of trustworthiness judgments and highlight the need to incorporate an intersectional approach to the study of social biases.

## Notes

1. We had access only to means and standard deviations already computed for each selected image. Participants recruited by the authors were 1,087. Gender: 552 females, 308 males and 227 did not report. Ethnicity: 516 White, 117 Asian, 74 Black, 72 biracial or multiracial, 57 Latinx, 18 other, and 233 did not report. The average age was 26.75 ( $SD = 10.54$ ).
2. Four counterbalanced IATs with women faces were presented, however only one task was correctly developed. The other three included errors in the coding that did not follow what was asked to participants to do in the instructions.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research received the EASP Seedcorn grant from the European Association of Social Psychology.

## Author's contributions

AV, EOM, MLH, PR and MB conceived the studies. AV, EOM, MLH performed the analysis. AV, EOM, MLH, PR and MB wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ORCID

A. Valmori  <http://orcid.org/0000-0003-3628-4204>  
 E. O. Meral  <http://orcid.org/0000-0002-6326-7840>  
 M. L. Hale  <http://orcid.org/0000-0002-7347-3445>  
 P. Rusconi  <http://orcid.org/0000-0002-0095-4061>  
 M. Brambilla  <http://orcid.org/0000-0002-4774-3309>

## References

- Barideaux, K., Jr., Crossby, A., & Crosby, D. (2021). Colorism and criminality: The effects of skin tone and crime type on judgments of guilt. *Applied Psychology in Criminal Justice*, 16, 181–199.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Brizi, A., & Mannetti, L. (2018). How competence and trustworthiness inferences made by voters in the centre of Italy predicts the North Italian political elections: The impact of the city size. *Rassegna Di Psicologia*, 33, 5–12. <https://doi.org/10.13133/1974-4854/16650>
- Calvert, G. A., Evans, G., & Pathak, A. (2022). Race, gender, and the U.S. presidency: A comparison of implicit and explicit biases in the electorate. *Behavioral Sciences*, 12(1), 17. <https://doi.org/10.3390/bs12010017>
- Carlsson, R., & Ågerström, J. (2016). A closer look at the discrimination outcomes in the IAT literature. *Scandinavian Journal of Psychology*, 57(4), 278–287. <https://doi.org/10.1111/sjop.12288>
- Carpenter, T. P., Pogacar, R., Pullig, C., Kouril, M., Aguilar, S., LaBouff, J., Isenberg, N., & Chakroff, A. (2019). Survey-software implicit association tests: A methodological and empirical analysis. *Behavior Research Methods*, 51(5), 2194–2208. <https://doi.org/10.3758/s13428-019-01293-3>
- Charbonneau, I., Robinson, K., Blais, C., & Fiset, D. (2020). Implicit race attitudes modulate visual information extraction for trustworthiness judgments. *PLoS One*, 15(9), e0239305. <https://doi.org/10.1371/journal.pone.0239305>
- Charlesworth, T. E. S., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. *Psychological Science*, 30(2), 174–192. <https://doi.org/10.1177/0956797618813087>
- Charlesworth, T. E. S., Caliskan, A., & Banaji, M. R. (2022). Historical representations of social groups across 200 years of word embeddings from Google books. *Proceedings of the National Academy of Sciences*, 119(28), e2121798119. <https://doi.org/10.1073/pnas.2121798119>
- Charnysh, V., Lucas, C., & Singh, P. (2015). The ties that bind: National identity salience and pro-social behavior toward the ethnic other. *Comparative Political Studies*, 48(3), 267–300. <https://doi.org/10.1177/0010414014543103>
- Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist*, 64(3), 170–180. <https://doi.org/10.1037/a0014564>
- Coles, S. M., & Pasek, J. (2020). Intersectional invisibility revisited: How group prototypes lead to the erasure and exclusion of Black women. *Translational Issues in Psychological Science*, 6(4), 314–324. <https://doi.org/10.1037/tps0000256>
- Collins, T., & Moyer, L. (2008). Gender, race, and intersectionality on the federal appellate bench. *Political Research Quarterly*, 61(2), 219–227. <https://doi.org/10.1177/1065912907306467>

- Deflem, M. (2022). Celebrity activism on racial justice during COVID-19: The death of George Floyd, the rise of Naomi Osaka, and the celebritization of race in pandemic times. *International Review of Sociology*, 32(1), 63–87. <https://doi.org/10.1080/03906701.2022.2052457>
- Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, 19(10), 978–980. <https://doi.org/10.1111/j.1467-9280.2008.02186.x>
- Dunivin, Z. O., Yan, H. Y., Ince, J., & Rojas, F. (2022). Black lives matter protests shift public discourse. *Proceedings of the National Academy of Sciences*, 119(10), e2117320119. <https://doi.org/10.1073/pnas.2117320119>
- Else-Quest, N. M., & Hyde, J. S. (2016). Intersectionality in quantitative psychological research. *Psychology of Women Quarterly*, 40(2), 155–170. <http://doi.org/10.1177/0361684316629797>
- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, 19(9), 1508–1519. <https://doi.org/10.1162/jocn.2007.19.9.1508>
- Franco, F. M., & Maass, A. (1999). Intentional control over prejudice: When the choice of the measure matters. *European Journal of Social Psychology*, 29(4), 469–477. [https://doi.org/10.1002/\(SICI\)1099-0992\(199906\)29:4<469::AID-EJSP938>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-0992(199906)29:4<469::AID-EJSP938>3.0.CO;2-S)
- Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). “Ain’t I a woman?”: Towards an intersectional approach to person perception and group-based harms. *Sex Roles*, 59(5–6), 392–403. <https://doi.org/10.1007/s11199-008-9505-4>
- Gomulya, D., Wong, E. M., Ormiston, M. E., & Boeker, W. (2017). The role of facial appearance on CEO selection after firm misconduct. *Journal of Applied Psychology*, 102(4), 617–635. <https://doi.org/10.1037/apl0000172>
- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the implicit association test can have societally large effects. *Journal of Personality and Social Psychology*, 108(4), 553–561. <https://doi.org/10.1037/pspa0000016>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. <http://doi.org/10.1037/0022-3514.85.2.197>
- Helman, E., Xie, S. Y., Ofosu, E. K., & Nespoli, G. A. (2018). Assessing the point at which averages are stable: A tool illustrated in the context of person perception [Preprint]. *PsyArXiv*. <https://doi.org/10.31234/osf.io/2n6jq>
- Hester, N., Payne, K., Brown-Iannuzzi, J., & Gray, K. (2020). On intersectionality: how complex patterns of discrimination can emerge from simple stereotypes. *Psychological Science*, 31(8), 1013–1024. <https://doi.org/10.1177/0956797620929979>
- Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, 14(6), 640–643. <https://doi.org/10.1046/j.0956-7976.2003.psci.1478.x>
- Kahn, K. B., & Davies, P. G. (2011). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Processes & Intergroup Relations*, 14(4), 569–580. <https://doi.org/10.1177/1368430210374609>
- Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, 111(5), 655–664. <https://doi.org/10.1037/pspa0000062>
- Kubota, J. T., Li, J., Bar-David, E., Banaji, M. R., & Phelps, E. A. (2013). The price of racial bias: Intergroup negotiations in the ultimatum game. *Psychological Science*, 24(12), 2498–2504. <https://doi.org/10.1177/0956797613496435>
- Livingston, R. W., & Pearce, N. A. (2009). The teddy-bear effect: Does having a baby face benefit Black chief executive officers? *Psychological Science*, 20(10), 1229–1236. <https://doi.org/10.1111/j.1467-9280.2009.02431.x>
- Livingston, R. W., Rosette, A. S., & Washington, E. F. (2012). Can an agentic Black woman get ahead? The impact of race and interpersonal dominance on perceptions of female leaders. *Psychological Science*, 23(4), 354–358. <https://doi.org/10.1177/0956797611428079>
- Lui, P. P., Parikh, K., Katedia, S., & Jouriles, E. N. (2022). Anti-Asian discrimination and antiracist bystander behaviors amid the COVID-19 outbreak. *Asian American Journal of Psychology*, 13(3), 295–304. <https://doi.org/10.1037/aap0000258>
- Lundberg, G. J. W., Neel, R., Lassetter, B., & Todd, A. R. (2018). Racial bias in implicit danger associations generalizes to older male targets. *PLoS One*, 13, e0197398. <https://doi.org/10.1371/journal.pone.0197398>
- 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(4), 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Mathews, T. J., & Johnson, G. S. (2015). Skin complexion in the twenty-first century: The impact of colorism on African American women. *Race, Gender & Class*, 22(1–2), 248–274.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37(5), 435–442. <https://doi.org/10.1006/jesp.2000.1470>
- McCoy, S. K., & Major, B. (2003). Group identification moderates emotional responses to perceived prejudice. *Personality and Social Psychology Bulletin*, 29(8), 1005–1017. <https://doi.org/10.1177/0146167203253466>
- Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, 34(2), 83–110. <https://doi.org/10.1007/s10919-009-0082-1>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of*

- Sciences*, 105(32), 11087–11092. <https://doi.org/10.1073/pnas.0805664105>
- Opie, T. R., & Phillips, K. W. (2015). Hair penalties: The negative influence of Afrocentric hair on ratings of Black women's dominance and professionalism. *Frontiers in Psychology*, 6, 1311. <https://doi.org/10.3389/fpsyg.2015.01311>.
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171–192. <https://doi.org/10.1037/a0032734>
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology*, 108(4), 562–571. <https://doi.org/10.1037/pspa0000023>
- Parent, M. C., DeBlaere, C., & Moradi, B. (2013). Approaches to research on intersectionality: Perspectives on gender, LGBT, and racial/ethnic identities. *Sex Roles*, 68(11–12), 639–645. <https://doi.org/10.1007/s11199-013-0283-2>
- Petsko, C. D., Rosette, A. S., & Bodenhausen, G. V. (2022). Through the looking glass: A lens-based account of intersectional stereotyping. *Journal of Personality and Social Psychology*, 123(4), 763–787. <https://doi.org/10.1037/pspi0000382>
- Platak, S. M., Krill, A. L., & Wilson, B. (2009). Implicit trustworthiness ratings of self-resembling faces activate brain centers involved in reward. *Neuropsychologia*, 47(1), 289–293. <https://doi.org/10.1016/j.neuropsychologia.2008.07.018>
- Rosenthal, L. (2016). Incorporating intersectionality into psychology: An opportunity to promote social justice and equity. *American Psychologist*, 71(6), 474–485. <https://doi.org/10.1037/a0040323>
- Rosette, A. S., Ponce de Leon, R., Koval, C. Z., & Harrison, D. A. (2018). Intersectionality: Connecting experiences of gender with race at work. *Research in Organizational Behavior*, 38, 1–22. <https://doi.org/10.1016/j.riob.2018.12.002>
- Sawyer, J., & Gampa, A. (2018). Implicit and explicit racial attitudes changed during Black lives matter. *Personality and Social Psychology Bulletin*, 44(7), 1039–1059. <https://doi.org/10.1177/0146167218757454>
- Schimmack, U. (2021). The implicit association test: A method in search of a construct. *Perspectives on Psychological Science*, 16(2), 396–414. <https://doi.org/10.1177/1745691619863798>
- 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. <https://doi.org/10.1016/j.jesp.2009.10.016>
- Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences*, 108(19), 7710–7715. <https://doi.org/10.1073/pnas.1014345108>
- Steffensmeier, D., Painter-Davis, N., & Ulmer, J. (2017). Intersectionality of race, ethnicity, gender, and age on criminal punishment. *Sociological Perspectives*, 60(4), 810–833. <https://doi.org/10.1177/0731121416679371>
- Stevens, F., & Shriver, E. (2024). The shame of implicit racial bias. *The Journal of Social Psychology*, 164(2), 258–272. <https://doi.org/10.1080/00224545.2022.2046538>
- strmic-pawl, H. V., Gonlin, V., & Garner, S. (2021). Color in context: Three angles on contemporary colorism. *Sociology of Race and Ethnicity*, 7(3), 289–303. <https://doi.org/10.1177/23326492211012532>
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127(1), 105–118. <https://doi.org/10.1016/j.cognition.2012.12.001>
- Vacchini, S., Fasoli, F., & Volpato, C. (2021). “You cannot get into my taxi!” Perceptions of a COVID-19-based rejection episode reported in the newspapers. *Journal of Language and Social Psychology*, 40(5–6), 677–689. <https://doi.org/10.1177/0261927X211043096>
- Walker, M., Jiang, F., Vetter, T., & Sczesny, S. (2011). Universals and cultural differences in forming personality trait judgments from faces. *Social Psychological and Personality Science*, 2(6), 609–617. <https://doi.org/10.1177/1948550611402519>
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, 26(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>