A blind spot for attractiveness discrimination

Jaeger, Bastian; Paolacci, Gabriele; Boegershausen, Johannes

A blind spot for attractiveness discrimination

Bastian Jaeger[1,2], Gabriele Paolacci[3], Johannes Boegershausen[3]

[1]Tilburg University, [2]Vrije Universiteit Amsterdam, [3]Erasmus University Rotterdam

Version: 11 November 2022
This paper is currently undergoing peer review. Comments are welcome.

Corresponding author:
Bastian Jaeger
Department of Social Psychology
Tilburg University
Room S406
Prof. Cobbenhagenlaan 225
5037DB Tilburg
The Netherlands.
E-mail: bxjaeger@gmail.com

Acknowledgments: The authors declare no conflicts of interest. The second and third author contributed equally to this paper. We are grateful to seminar participants at the VU Amsterdam, Mario Pandelaere, Peeter Verlegh, and Karl Aquino for their feedback on this project.

Open practices statement: Studies 3-7 and Study S1 were preregistered. All data, analysis scripts, materials, and preregistration documents are available at the Open Science Framework (https://osf.io/v2j3p/). We report how our sample sizes were determined and mention all data exclusions and measures for each study.

Abstract

Discrimination remains a key challenge for social equity. There is widespread agreement that discrimination is unfair and should be punished. A prerequisite for this is that instances of discrimination are detected. Yet, some types of discrimination may be less apparent than others. Across seven studies (N = 3,486, five preregistered), we find that attractiveness discrimination often goes undetected compared to more prototypical types of discrimination (i.e., gender and race discrimination). This blind spot does not emerge because people perceive attractiveness discrimination to be unproblematic or desirable. Rather, our findings suggest that people's ability to detect discrimination is bounded. People only focus on a few salient dimensions, such as gender and race, when scrutinizing decision outcomes (e.g., hiring or sentencing decisions) for bias. Consistent with this account, two interventions that increased the salience of attractiveness increased the detection of attractiveness discrimination, but also decreased the detection of gender and race discrimination.

Keywords: attractiveness; bias; discrimination; debiasing

Discrimination has harmful consequences for individuals, organizations, and society. Decades of research have shown that people are often treated unjustly because of their gender, race, sexual orientation, religious affiliation, or other personal characteristics [1]. Discrimination is widespread across several diverse domains, including employee selection, medical treatment, and criminal sentencing [2-4]. Most people agree that discrimination is unfair, and reducing discrimination is a prominent concern for politicians and activists across many societies. To combat discrimination, however, one first needs to *detect* it [5-8]. Whereas some types of discrimination are easily detected by observers, others may be less obvious and not readily detected [9]. A failure to detect discrimination can lead to fewer sanctions for those who discriminate and less support for the victims of discrimination. More generally, public responses to some types of discrimination may be lacking or inadequate simply because they are not detected. In this paper, we document a blind spot for attractiveness discrimination. We find that comparable instances of discrimination (e.g., biased selections of job candidates) are less likely to be detected when decision-makers discriminate based on attractiveness (vs. race and gender).

Converging evidence from social psychology, political science, economics, and biology indicates that people are favorably biased toward attractive people (often referred to as the "what is beautiful is good stereotype," the "beauty premium," or "lookism" [10,11,12]). Attractive individuals are seen as more trustworthy, competent, and socially skilled [13]. For many domains, partiality towards attractive individuals appears to be unsubstantiated or greatly exaggerated, as associations between physical attractiveness and various measures of personality and ability are small to non-existent [14-16]. This can lead to instances of discrimination against unattractive individuals. Suggestively, attractive people earn higher wages [17-19] and are more likely to be promoted [20] and selected for leadership positions [21]. Field experiments in which the attractiveness of candidates was manipulated while keeping all other information constant found higher callback rates for attractive candidates [22,23]. Outside the employment domain, attractive politicians receive more votes [24], attractive defendants are advantaged in the courtroom [25], and attractive teachers receive more positive course evaluations [26]. Relatedly, people generally act more prosocially toward attractive individuals [27].

Even though attractiveness discrimination is widespread, it has received relatively little attention in public debate. There are few social movements that aim to curb attractiveness discrimination and few laws and regulations specifically targeted at providing protection against

attractiveness discrimination [19]. This lack of sociopolitical salience was also noted in a recent New York Times article, which raised the puzzling question of why "prejudice against the unattractive […] gets almost no attention and sparks little outrage" [28].

We draw on social-cognitive theories to test one explanation for the mismatch between the prevalence of attractiveness discrimination and the lack of concern for it. Attractiveness discrimination is less likely to be detected compared to more prototypical types of discrimination [29-31]. In almost all situations, it is not necessarily apparent whether specific decisions (e.g., rejecting job candidates, promoting employees, convicting defendants) result from a discriminatory decision process. When judging whether individual outcomes are due to discrimination, people may compare them to the mental prototypes of discrimination. For instance, the same decision is more likely to be attributed to discrimination if it involves more prototypical perpetrators [32] and victims [33] of discrimination (e.g., a male recruiter rejecting a female job candidate). Context also influences attributions. For example, rejecting a female candidate was seen as more discriminatory if the job in question required masculine (vs. feminine) traits, whereas the opposite pattern was observed for male candidates [34]. Thus, when outcomes fit with people's mental image of discrimination, such as when the historically disadvantaged experience negative outcomes because of decisions by the historically advantaged, they are more likely to be perceived as discriminatory.

Previous studies mostly focused on attributions to discrimination of ambiguous decisions. We build on these insights and test if even clear instances of discrimination (e.g., hiring decisions that clearly favor one type of applicant) can go undetected if the type of discrimination is less prototypical. Gender and race are salient categories in person perception [35] and gender and race discrimination are viewed as the most salient forms types of discrimination [30]. Thus, we hypothesized that although people readily detect instances of race and gender discrimination (i.e., more prototypical types of discrimination), this would not be true of attractiveness discrimination (i.e., a less prototypical type of discrimination).

We present the results of seven studies (total N = 3,486, five preregistered). Figure 1 illustrates the basic experimental paradigm. Participants evaluate the outcomes of a selection process. They are presented with the consideration set of the decision-maker, containing 24 job candidates. The candidates vary in gender (12 male, 12 female), race (12 Black, 12 White), and attractiveness (12 attractive, 12 unattractive), each represented by a facial photograph (see the

Methods section for more details on stimulus selection and validation). Participants are told that all candidates are highly qualified and fulfill all requirements of the company. Then, participants are shown the candidates that were hired by the company and evaluate the fairness of the selection process. We manipulate whether selection outcomes are strongly biased in favor of one type of candidate, suggesting the presence of discriminatory hiring practices. Participants see selections that are attractiveness-biased (only attractive candidates selected), selections that are gender-biased (only male candidates selected), selections that are race-biased (only White candidates selected), and selections that are unbiased (candidates are balanced on attractiveness, gender, and race). In each condition, only one type of selection bias was implemented. For example, in the gender bias condition, all selected candidates were male, but they were otherwise balanced on attractiveness and race.

Our primary goal was to test whether people are less likely to detect instances of attractiveness discrimination than instances of gender or race discrimination. We test this hypothesis in different contexts (hiring, legal sentencing), using different dependent variables (i.e., direct and indirect measures of detection), different study designs (i.e., between-subjects and within-subjects designs), and different populations (undergraduate students, representative samples of U. S. Americans). Studies 1-5 examine the robustness and generalizability of people's blind spot for attractiveness discrimination, ruling out several alternative explanations. Study 6 investigates whether the blind spot emerges because people, when scrutinizing decision outcomes for bias, primarily pay attention to types of discrimination that are more prototypical than attractiveness. Finally, Study 7 tests the effectiveness (and the potential for unintended consequences) of two interventions designed to increase the detectability of attractiveness discrimination.

We report the results of frequentist and Bayesian analyses for most tests (see Supplemental Materials for an explanation of how Bayes Factors were approximated [36]). Bayes Factors (BFs) always indicate the strength of support for the favored hypothesis (i.e., BF_{10} [BF_{01}] when evidence favors the alternative [null] hypothesis). All data, materials, preregistration documents, and analysis scripts are available at the Open Science Framework (https://osf.io/v2j3p/).

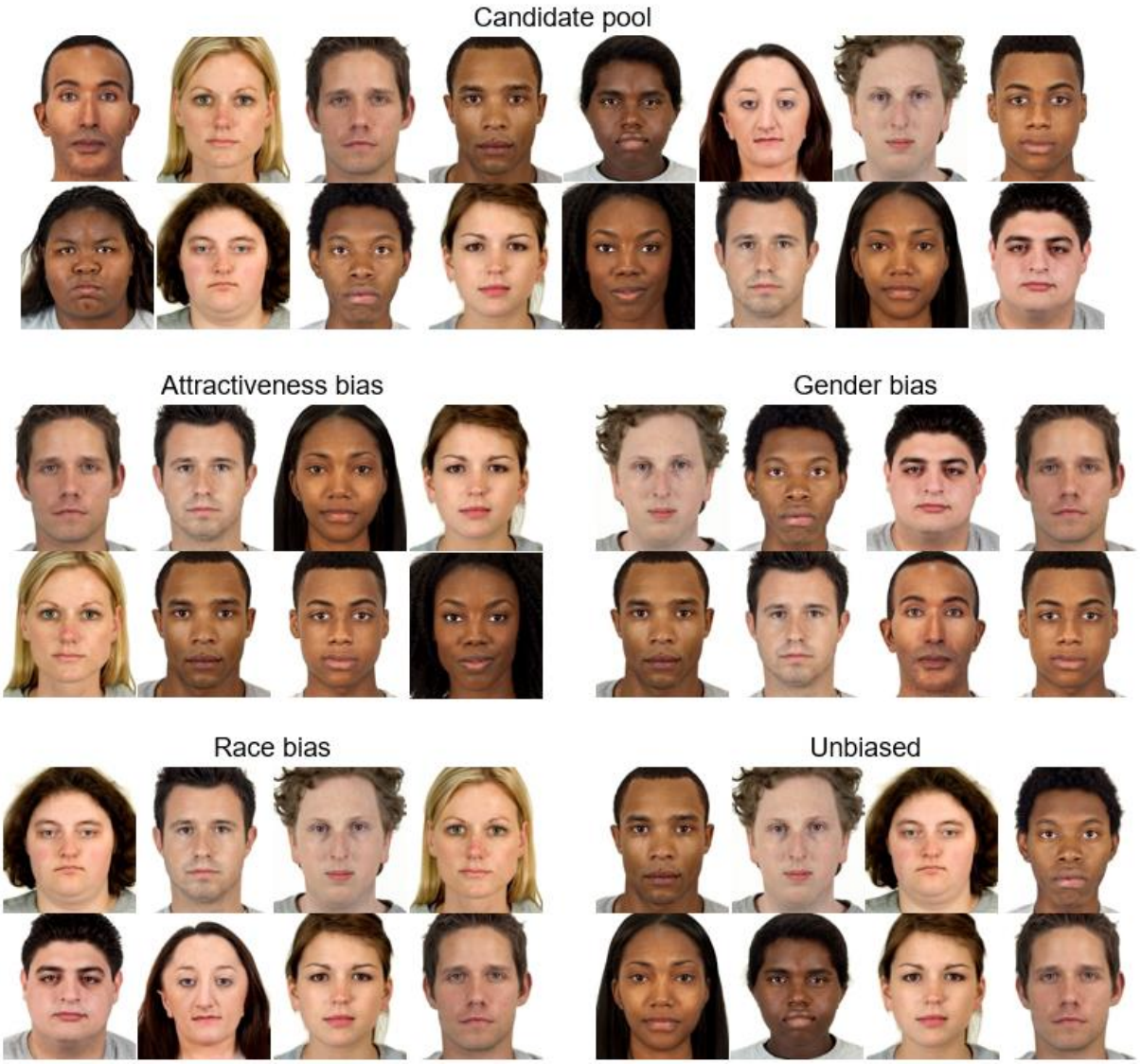

Figure 1. Exemplary selection of candidates from Study 2 displaying an attractiveness bias (only attractive candidates selected), gender bias (only male candidates selected), race bias (only White candidates selected), or no bias (balanced on attractiveness, gender, and race). Note that participants always first saw the initial pool of 24 considered candidates.

Results

Study 1

Studies 1-4 were designed to test whether people are less likely to detect instances of attractiveness discrimination (vs. gender or race discrimination). In Study 1 (n = 599), we measured discrimination detection in a free-response format to test if people *spontaneously* reference discriminatory selection practices when they see attractiveness-, gender-, or race-biased selections of candidates [30,31]. Participants first saw the initial pool of 24 candidates

varying in attractiveness, gender, and race, and then one group of 8 selected candidates. Type of discrimination was manipulated in a between-subjects design and participants saw either an attractiveness-, gender-, or race-biased selection. We asked participants to write down anything that "stood out" about the hiring process. Three coders who were blind to the hypotheses coded whether participants mentioned that selections were based on candidates' attractiveness, gender, or race. Whereas 65.00% of participants perceived gender discrimination in the gender-biased condition and 72.00% perceived racial discrimination in the race-biased condition, only 23.12% of participants in the attractiveness-biased condition referred to candidates' attractiveness as a potential selection criterion. Regressing a binary variable indicating whether the correct type of discrimination was mentioned on a dummy variable indicating which condition participants were in confirmed that attractiveness discrimination was referenced less often than gender discrimination, β = -1.82, SE = 0.22, z = 8.12, p < .001, BF_{10} > 1000, and race discrimination, β = -2.15, SE = 0.23, z = 9.32, p < .001, BF_{10} > 1000. There was no significant difference in detection rates for gender vs. race discrimination, β = 0.33, SE = 0.22, z = 1.51, p = .132, BF_{01} = 6.41. Thus, the results were in line with the hypothesis that attractiveness discrimination goes undetected more often than gender or race discrimination. Results were similar when examining whether participants made any reference to attractiveness, gender, or race in their descriptions (see Supplementary Information). This yielded very similar results.

Study 2

Rather than counting spontaneous references to discrimination, Study 2 (n = 199) measured the extent to which participants perceived discrimination by asking them to judge the fairness of the selection process [33,37]. Using our basic experimental paradigm, participants saw a pool of 24 qualified job candidates varying in attractiveness, gender, and race. They then saw 12 groups of 8 candidates that were hired for the job. We varied whether selections were attractiveness-biased, gender-biased, race-biased, or unbiased (within-subject manipulation, three trials per condition). Our analysis focused on differences in fairness ratings between the unbiased condition and each of the three bias conditions to test if participants detected that candidates were selected based on certain attributes. Higher fairness ratings suggest a lower sensitivity to discrimination. Compared to the unbiased condition (M = 5.68, SD = 1.18), participants rated candidate selections as less fair in the attractiveness bias condition (M = 5.39, SD = 1.27), β = -0.29, t(197) = -3.49, SE = 0.08, p < .001, BF_{01} = 2.29, in the gender bias condition (M = 2.80, SD

= 1.59), β = -2.88, *SE* = 0.12, *t*(197) = -23.81, $p < .001$, $BF_{01} > 1000$, and in the race bias condition (*M* = 3.09, *SD* = 1.66), β = 2.59, *SE* = 0.12, *t*(197) = -21.35, $p < .001$, $BF_{10} > 1000$. However, there were large differences in how different selection biases affected fairness perceptions (see Figure 2). The decrease in perceived fairness resulting from biased (relative to unbiased) selection was markedly smaller for attractiveness bias compared to gender bias, z = 15.70, $p < .001$, and race bias, z = 17.72, $p < .001$.

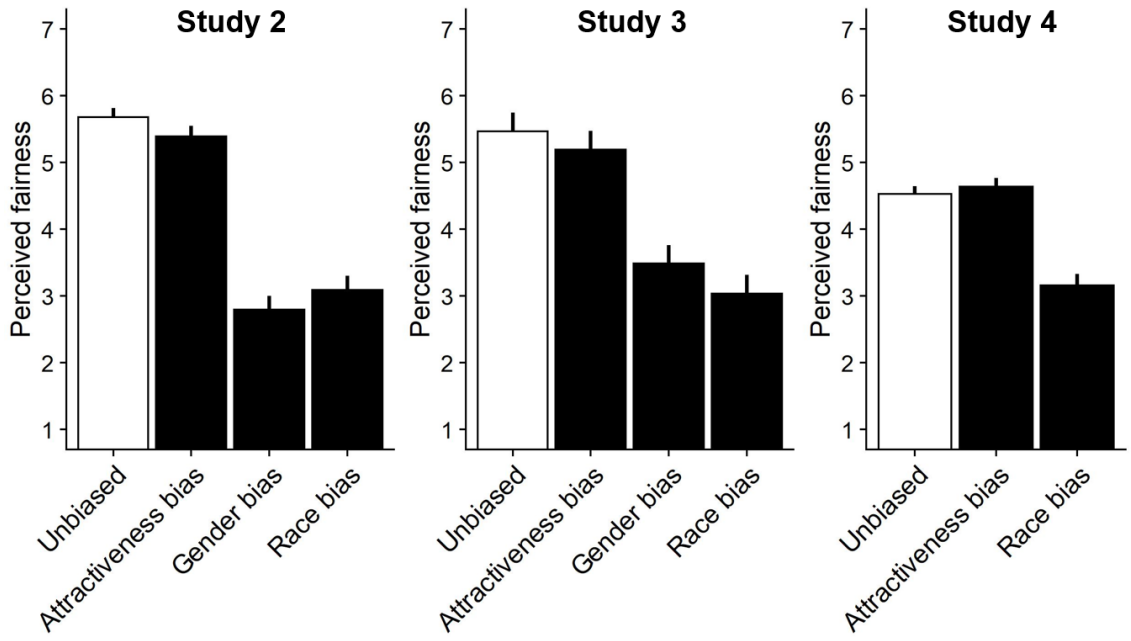

Figure 2. Fairness ratings of unbiased, attractiveness-biased, gender-biased, and race-biased selections in Study 2, Study 3, and Study 4. Higher fairness ratings indicate lower sensitivity to discrimination. In Studies 2 and 3, participants judged the outcomes of hiring procedures. In Study 4, participants judged the outcomes of criminal sentencing procedures. Error bars indicate 95% confidence intervals.

Study 3

Study 3 (*n* = 402, preregistered: https://aspredicted.org/KKG_JUA) tested the robustness of the results in Study 2 by manipulating the type of bias in a between-subjects design. We again found that compared to the unbiased condition (*M* = 5.47, *SD* = 0.42), participants rated candidate selections as less fair in the gender bias condition (*M* = 3.49, *SD* = 0.46), β = -1.98, *SE* = 0.20, *t*(398) = -9.83, $p < .001$, $BF_{10} > 1000$, and in the race bias condition (*M* = 3.04, *SD* = 0.37), β = -2.43, *SE* = 0.20, *t*(398) = -12.03, $p < .001$, $BF_{10} > 1000$ (see Figure 2). The difference in fairness ratings between the unbiased condition and the attractiveness-biased condition (*M* =

5.19, *SD* = 0.46) was not significant, β = -0.28, *SE* = 0.20, *t*(398) = -1.37, *p* = .173, BF_{01} =13.23. We again found that the decrease in perceived fairness resulting from biased selection (vs. unbiased selection) was much smaller for attractiveness discrimination compared to gender discrimination, *z* = 15.70, p < .001, and race discrimination, *z* = 17.72, *p* < .001. Note that a between-subjects design makes it impossible for people to compare perceived biases across conditions. Therefore, these results speak against an alternative account for the results of Study 2, i.e., that participants are equally likely to recognize different types of bias but assign lower fairness ratings to selections that are seen as relatively less tolerable (gender- and race-biased selections as opposed to attractiveness-biased). The next two studies address this possibility more directly.

Study 4

In Study 4 (*n* = 300, preregistered: https://aspredicted.org/QTC_YSG), we tested whether our results generalize beyond a hiring context and to observed decisions of whom to penalize (vs. reward). Participants considered 20 male defendants varying in attractiveness and race. They were told that all defendants had committed the same type of crime and that the evidence regarding their guilt was similarly ambivalent. Participants then considered 9 groups of 8 defendants that were found guilty. In a within-subject design that included three trials per condition, we varied whether sentencing decisions were attractiveness-biased (only unattractive defendants were found guilty), race-biased (only Black defendants were found guilty), or unbiased. Compared to the unbiased condition (*M* = 4.53, *SD* = 1.26), participants rated verdicts as less fair in the race bias condition (*M* = 3.16, *SD* = 1.62), β = -1.37, *SE* = 0.09, *t*(299) = -15.06, *p* < .001, BF_{10} > 1000, but as slightly *more* fair in the attractiveness bias condition (*M* = 4.64, *SD* = 1.28), β = 0.11, *SE* = 0.04, *t*(299) = 2.66, *p* = .008, BF_{01} = 12.15 (see Figure 2). Again, the difference in perceived fairness between a biased and an unbiased decision was much smaller for attractiveness bias than for race bias, *z* = 12.58, *p* < .001. The results of Study 4 thus replicate our effect that observers are less sensitive to attractiveness discrimination also in a setting where it manifests itself as selecting less (vs. more) attractive people for adverse (vs. favorable) outcomes.

We also explored whether fairness ratings of different types of discrimination are moderated by individual differences in two traits that have been linked to how people react to (potential) instances of discrimination, namely social dominance orientation SDO [38] and

perceived everyday discrimination PED [39]. First, we regressed fairness ratings on condition (attractiveness bias, race bias, unbiased), SDO, and their interaction. We found a significant interaction effect with SDO when comparing the unbiased and race bias conditions, $\beta = 0.47$, $SE = 0.09$, $t(298) = 5.38$, $p < .001$, $BF_{10} = 5474$. Race-biased verdicts were rated as less fair than unbiased verdicts, and this difference was more pronounced among participants who scored lower on SDO (i.e., who endorsed egalitarian values more strongly). However, we did not find a significant interaction effect with SDO when comparing the unbiased and attractiveness bias conditions, $\beta = 0.02$, $SE = 0.04$, $t(298) = 0.45$, $p = .65$, $BF_{01} = 365.7$.

Results were similar for individual differences in perceived everyday discrimination. Race-biased verdicts were rated as less fair than unbiased verdicts and this difference was more pronounced among participants who scored higher on PED, $\beta = -0.32$, $SE = 0.09$, $t(298) = -3.61$, $p < .001$, $BF_{10} = 3.21$. However, the difference in fairness ratings between attractiveness-biased and unbiased verdicts was not significantly moderated by individual differences in PED, $\beta = -0.03$, $SE = 0.04$, $t(298) = -0.71$, $p = .48$, $BF_{01} = 313.8$. These results lend further support to our proposition that attractiveness discrimination receives less attention compared to other biases because it is less likely to be detected, and not because people detect this bias but perceive it as less egregious than other biases. If participants detected attractiveness discrimination, then fairness ratings of attractiveness-biased outcomes should be lower among participants who tend to show stronger negative reactions to discrimination. This was not the case, however, as individual differences in egalitarianism and perceived everyday discrimination only explained the extent to which people reacted negatively to instances of race discrimination, but not to instances of attractiveness discrimination. Study 5 further examines the possibility that attractiveness discrimination is detected but not deemed as unfair compared to other biases.

Study 5

Similar to Study 3, in Study 5 ($n = 720$, preregistered: https://aspredicted.org/QSG_AJT) we manipulated the selection bias (attractiveness vs. gender vs. race) in a between-subjects design, and asked participants to rate the fairness of the selection procedure. Additionally, we manipulated whether the bias was made explicit to participants. In the *control condition*, participants only saw the selected candidates (as in previous studies). In the *explicit discrimination condition*, participants were told that the recruiter "selected attractive/White/male candidates and rejected unattractive/Black/female candidates." For example, in the attractiveness

bias condition, participants were told that the recruiter decided to hire the most attractive candidates and reject the least attractive candidates. We were particularly interested in how this manipulation would affect participants' fairness perceptions of attractiveness-biased selections. If people view attractiveness discrimination as unfair but fail to detect it, then making the bias explicit should lower fairness ratings. Alternatively, if participants detect the selection bias, but do not perceive it as unfair, then making the bias explicit should not affect their ratings. Results supported the former hypothesis. As predicted, regressing fairness ratings on the type of bias, explicitness, and their interaction, revealed a significant interaction effect, $F(2, 714) = 23.78$, $p < .001$. Making the selection bias explicit significantly decreased fairness ratings for gender-biased selections, $\beta = -0.86$, $SE = 0.20$, $t(238) = -4.21$, $p < .001$, $BF_{01} = 355.9$, race-biased selections, $\beta = -0.90$, $SE = 0.19$, $t(239) = -4.81$, $p < .001$, $BF_{10} > 1000$, and critically, for attractiveness-biased selections, $\beta = -2.59$, $SE = 0.22$, $t(237) = -11.97$, $p < .001$, $BF_{10} > 1000$ (see Figure 3). Importantly, the reduction in perceived fairness was much more pronounced for attractiveness discrimination compared to gender discrimination, $\beta = -1.73$, $SE = 0.30$, $t(714) = 6.03$, $p < .001$, $BF_{10} > 1000$, and race discrimination, $\beta = -1.69$, $SE = 0.29$, $t(714) = 5.92$, $p < .001$, $BF_{10} > 1000$. These results suggest that people do object to attractiveness discrimination when they are aware that attractiveness discrimination occurred, but they are unlikely to detect it and react to it by themselves. Our final two studies suggest that this may happen because attractiveness competes for people's attention with more prototypical types of discrimination.

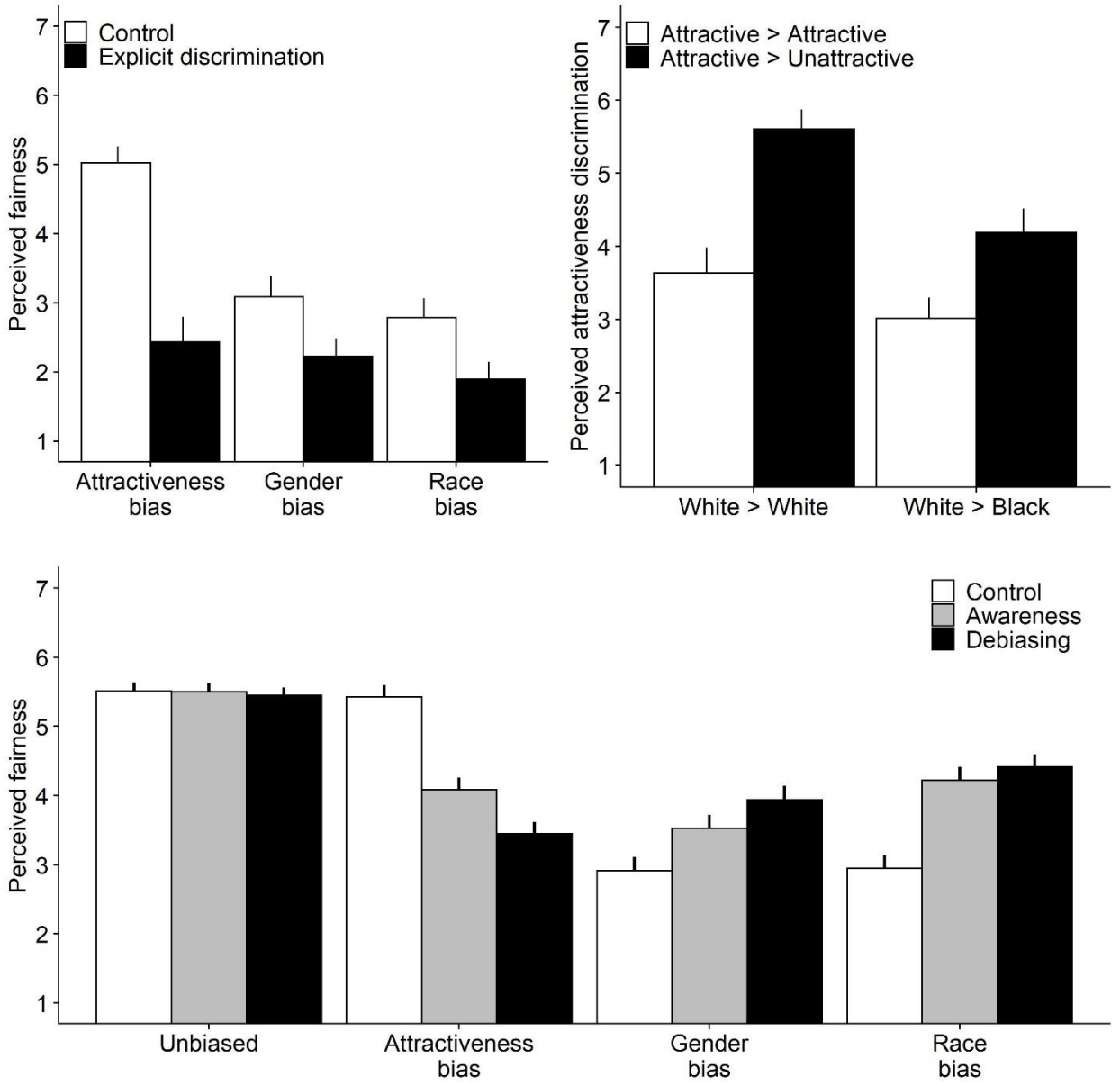

Figure 3. Fairness ratings of attractiveness-biased, gender-biased, and race-biased selections when the recruiter's bias was made and was not made explicit (Study 5). Perceived attractiveness discrimination as a function of whether the attractive White candidate was selected over an attractive vs. unattractive, White vs. Black candidate (Study 6). Fairness ratings with no intervention, with an intervention that made participants aware of attractiveness discrimination, and with an intervention that made participants aware of it and instructed them to pay attention to candidates' attractiveness (Study 7). Error bars indicate 95% confidence intervals. Higher fairness ratings (Study 5 and Study 7) indicate lower sensitivity to discrimination, whereas higher perceived attractiveness discrimination ratings indicate higher sensitivity to discrimination.

Study 6

Study 6 (n = 505, preregistered: https://aspredicted.org/9FB_3NJ) examined whether people's blind spot for attractiveness discrimination can be partly explained by their focus on more prototypical types of discrimination. In other words, people's alertness to race and gender as widespread types of discrimination might make it less likely that people attend to and detect attractiveness discrimination. We tested this possibility by varying whether a potential instance of attractiveness discrimination (i.e., an attractive candidate being chosen over an unattractive candidate) could also be explained by race discrimination.

Participants considered two male job candidates –of which only one was selected– and rated the extent to which the recruiter's decision was due to attractiveness discrimination. We measured perceived attractiveness discrimination with an item that was embedded with other filler items (see supplementals). The selected candidate was White and attractive. Crucially, in a 2 × 2 between-subjects design, we varied whether the rejected candidate was attractive or unattractive (thus varying the potential for attributing the recruiter's decision to attractiveness discrimination) and whether he was Black or White (thus varying the potential for attributing the recruiter's decision to race discrimination). We reasoned that if the detectability of attractiveness bias is caused by people's focus on more prototypical types of discrimination, then participants should be less likely to attribute a hiring decision to attractiveness discrimination (i.e., a decision that favored a candidate because he is relatively more attractive) when the decision could also be explained by race discrimination (i.e., a decision that favored a candidate because he is White rather than Black).

We regressed perceptions of attractiveness discrimination on the attractiveness of the rejected candidate, their race, and the interaction between these two factors. We found a significant interaction effect, β = -0.80, SE = 0.32, $t(501)$ = -2.52, p = .012, BF_{10} = 1.07 (see Figure 3). Not surprisingly, participants were more likely to perceive attractiveness discrimination when an attractive candidate was favored over an unattractive candidate compared to an attractive candidate being favored over another attractive candidate, β = 1.57, SE = 0.16, $t(501)$ = 9.94, p < .001, BF_{10} > 1000. Crucially, this effect was more pronounced when the White candidate was chosen over another White candidate, β = 1.97, SE = 0.22, $t(251)$ = 8.94, p < .001, BF_{10} > 1000, compared to when the White candidate was chosen over the Black candidate, β = 1.17, SE = 0.23, $t(250)$ = 5.18, p < .001, BF_{10} > 1000. In other words, when a recruiting decision could be

explained by race discrimination, participants were less likely to see it as affected by attractiveness discrimination.

Study 7

In Study 7 (n = 761, preregistered: https://aspredicted.org/9FJ_C5V), we tested the effectiveness of two interventions aimed at diminishing people's blind spot for attractiveness discrimination. We also explored whether the interventions would affect the detection of other types of discrimination. As in Study 2, participants rated the fairness of attractiveness-biased, gender-biased, race-biased, and unbiased selections (within-subject manipulation). We tested the effect of two interventions differing in strength. In the *awareness condition*, participants read a news article about beauty biases before viewing the selected candidates and indicating their fairness ratings. The article summarized research on attractiveness discrimination in the workplace and in a legal context, and made the argument that attractiveness discrimination can lead to unfair inequalities, just like gender, race, and other types of discrimination. In the *debiasing condition*, participants read the same article but were also instructed to consider whether candidate attractiveness might have influenced the recruiter's decision. They were told that attractiveness discrimination often goes unnoticed and that they should pay special attention to candidates' attractiveness when judging the fairness of the selection process. In the *control condition*, participants read a news article about the horseshoe crab.

First, we tested whether the two interventions influenced participants' fairness ratings of the attractiveness-biased selection. Compared to the control condition, participants viewed attractiveness-biased selections as less fair after the awareness intervention, β = -1.34, SE = 0.08, $t(1510) = -17.43$, $p < .001$, $BF_{10} > 1000$, and after the debiasing intervention, β = -1.97, SE = 0.08, $t(1534) = -26.29$, $p < .001$, $BF_{10} > 1000$, (see Figure 3). Perceived fairness was significantly lower in the debiasing condition than in the awareness condition, β = -0.64, SE = 0.09, $t(1516) = -7.46$, $p < .001$, $BF_{10} > 1000$.

Next, we tested whether the interventions selectively affected participants' reactions towards attractiveness-biased selections or whether they simply lowered fairness ratings in general, that is, also in cases where there was no apparent selection bias. When participants rated unbiased selections of candidates, there were no significant differences in fairness ratings between the control condition and the two intervention conditions (awareness vs. control: β = -

0.01, *SE* = 0.06, *t*(1510) = -0.17, *p* = .864, BF_{01} = 38.32; debiasing vs. control: β = -0.07, *SE* = 0.06, *t*(1534) = -1.10, *p* = .272, BF_{01} = 21.43).

Finally, we examined the effect of the interventions on participants' ratings of the gender- and race-biased selections. Three patterns of results were plausible. If the interventions only affect participants' detection of attractiveness discrimination, without affecting the detection of other types of discrimination, then we would expect no effects of the interventions. Alternatively, the interventions may have a spillover effect in that they make participants more sensitive to all types of discrimination. In this case, we would expect the interventions to also reduce fairness ratings in the gender bias and race bias conditions. Finally, the opposite scenario is also plausible, in that people's capacity to detect discrimination may be bounded. That is, the interventions may divert participants' attention toward attractiveness but away from race and gender. The current results support this last hypothesis. Compared to the control condition, participants rated gender-biased selections as *more* fair after the awareness intervention, β = 0.61, *SE* = 0.09, *t*(1510) = 6.65, *p* < .001, BF_{10} > 1000, and after the debiasing intervention, β = 1.03, *SE* = 0.09, *t*(1534) = 11.56, *p* < .001, BF_{10} > 1000. Additionally, compared to the control condition, participants rated race-biased selections as *more* fair after the awareness intervention, β = 1.27, *SE* = 0.09, *t*(1510) = 14.43, *p* < .001, BF_{10} > 1000, and after the debiasing intervention, β = 1.46, *SE* = 0.08, *t*(1534) = 17.84, *p* < .001, BF_{10} > 1000.

These findings suggest that both interventions increase the detection of attractiveness discrimination. The interventions did not simply cause participants to lower their fairness ratings across the board. We observed no effect on fairness ratings when there was no apparent bias. However, it seems that directing participants' attention toward attractiveness made them less likely to detect gender and race discrimination. This finding suggests that people's capacity to "scan for" biases might be limited. Although people spontaneously attend to gender and race discrimination, directing their attention to other sources of bias (e.g., attractiveness) reduces their sensitivity to instances of gender and race discrimination.

Discussion

We investigated whether people show a blind spot for attractiveness discrimination. People are more likely to perceive instances of discrimination that are more prototypical. For example, outcomes are more likely to be attributed to discrimination if they involve more prototypical perpetrators and victims of discrimination [34,40,41]. Here, we tested whether mental

prototypes can also influence which *types* of discrimination are detected. Across seven studies, we found that people are unlikely to detect attractiveness discrimination, particularly compared to race and gender discrimination.

Our studies yielded five core insights. First, participants' blind spot for attractiveness discrimination was robust and widespread. We found similar patterns of results when examining judgments of different outcomes (job applicant selections and legal sentencing), with different measures of discrimination perceptions (attributions to attractiveness discrimination, fairness ratings, open-ended descriptions), different manifestations of attractiveness discrimination (favoring attractive individuals and disfavoring unattractive individuals) with participants from different (Western) countries, and when manipulating the type of discrimination using within-subject or between-subjects designs. We found no evidence that individual difference measures that are typically related to how people perceive and evaluate instances of potential discrimination moderate the focal effect. Attractiveness discrimination went undetected even among participants who reported seeing more instances of discrimination in everyday life and who hold stronger egalitarian values.

Second, even obvious instances of attractiveness discrimination often go undetected. We selected the most attractive and least attractive individuals from a large and diverse stimulus set. Analyzing attractiveness ratings across the studies showed that around 95% of participants perceived the attractive targets as more attractive than the unattractive targets, and manipulation checks yielded large differences between the groups (see Supplementary Information, Figure S2). Moreover, participants always viewed selections in which *all* selected individuals were highly attractive. This presentation format and the absence of any additional information that could be used to "explain away" biases have been shown to aid in detecting bias [42]. Thus, attractiveness discrimination often remained undetected even though our studies facilitated detection.

Third, our results cannot be explained by the fact that participants view attractiveness-biased selection as unproblematic, fair, or not worth mentioning. When we made explicit that the recruiter selected based on candidates' attractiveness, hiring outcomes were rated as much less fair (Study 5). These findings support the interpretation that people view attractiveness discrimination as unfair but fail to detect it on their own.

Fourth, we found that participants' likelihood of attributing a hiring decision to attractiveness discrimination depended on whether the decision could also be attributed to race discrimination (Study 4). This finding is in line with the notion that people fail to detect attractiveness discrimination because they only pay attention to more prototypical types of discrimination (e.g., gender and race).

Fifth, interventions can overcome people's blind spot for attractiveness discrimination. The results of Study 7 suggest that raising awareness of attractiveness discrimination helps people detect instances of attractiveness discrimination. The effect was even larger when we also instructed participants to pay attention to attractiveness discrimination in the task. However, the interventions made participants *less* likely to detect instances of gender and race discrimination, suggesting that people pay attention to a limited number of dimensions when evaluating whether discrimination occurred. People may spontaneously focus on prototypical dimensions of discrimination, such as gender and race, which explains their blind spot for attractiveness discrimination. However, when their focus shifts to other dimensions due to intrinsic or extrinsic causes, they are less likely to detect gender and race discrimination. In other words, people have a bounded capacity to detect discrimination, in that they only pay attention to a few potential types of discrimination.

If people's blind spot for attractiveness discrimination emerges because it does not fit with prototypes of discriminatory acts, similar blind spots should be observed for other non-prototypical types of discrimination (e.g., weight bias). A prototype account would also predict that attractiveness discrimination is more readily perceived in domains where attractiveness discrimination is more common and, therefore, salient, such as when selecting sales personnel. These are important predictions for future studies to address. Future studies should also test whether the current results replicate when using more diverse samples of participants and targets. In the present studies, we primarily relied on participants and target stimuli from the United States.

More work is needed to explore how blind spots for discrimination can be mitigated. Making participants aware of attractiveness discrimination did increase detection of attractiveness discrimination, but it also had the unintended side effect of reducing the detection of gender and race discrimination. Raising awareness of multiple biases [43] or making people spend more time and cognitive effort on checking for discrimination based on various

dimensions may increase the detection of attractiveness discrimination without decreasing the detection of other types of discrimination. Other characteristics of the decision-making context may also influence the probability that attractiveness discrimination is detected. For example, people are less likely to recognize discrimination when it favors (vs. disfavors) certain types of individuals [7].

The present studies primarily focused on the spontaneous detection of different types of discrimination. Although we found some evidence that fairness perceptions of attractiveness-, gender-, and race-biased outcomes are very similar when the decision-maker's discriminatory intent was clear (Study 3), more work is needed to understand how people judge different types of explicit or blatant discrimination. This may depend on the characteristics of the decision-making context (e.g., attractiveness discrimination in recruiting may be seen as more acceptable when performance is directly linked to a person's appearance) or the decision-maker.

Methods

The studies were approved by the Institutional Review Boards of Tilburg University (Study 2) and the Erasmus Research Institute in Management (Erasmus University Rotterdam; Study 1, Studies 3-7, Study S1). All experiments complied with relevant ethical regulations, and informed consent was obtained from all participants. No participant completed more than one study reported here. All statistical tests were two-sided.

Study 1

Participants. We recruited 599 Prolific workers from the United States (M_{age} = 33.12 years, SD_{age} = 9.03; 35.06% female, 63.61% male, 1.34% non-binary; 71.79% White, 21.37% Black/African American, 1.66% American Indian/Alaska Native, 4.67% Asian, 0.16% Native Hawaiian/Pacific Islander, 3.84% Hispanic/Latinx, 1.00% other, 0% no response).

Stimuli. We selected 24 images from the Chicago Face Database [44]. All targets were instructed to adopt a neutral facial expression and targets were photographed from a standardized distance in front of a white background. Targets consented to their photographs being used in scientific articles. We selected targets that were balanced on gender (male or female), race (Black or White), and attractiveness (low or high). To select targets that are perceived as particularly attractive and unattractive, we relied on the validation data of the database. This data set contains self-reported data by targets (e.g., gender, race) and ratings of all targets on various dimensions, including attractiveness. Targets were rated by large and demographically diverse sample of 1,087 raters (M_{age} = 26.75, SD_{age} = 10.54; 47.47% White, 10.76% Asian, 6.81% Black, 6.62% biracial or multiracial, 5.24% Latino, 1.66% other, and 21.44% did not report; and 50.78% female, 28.33% male, and 20.88% did not report). Each target's attractiveness was evaluated by an average of 54 independent raters (*min* = 23, *max* = 97), which is sufficient to yield reliable average ratings [45]. Raters judged targets on a seven-point scale and average scores per target were created (see Ma and colleagues[44]). For each gender-by-race combination, we selected the three targets with the lowest average attractiveness (M_{age} = 1.91, SD_{age} = 0.20 for the 12 unattractive targets) and the three targets with the highest average attractiveness (M_{age} = 4.86, SD_{age} = 0.38 for the 12 attractive targets).

Procedure. Participants saw the 24 targets varying in gender, race, and attractiveness and were told to imagine that these individuals applied for a job with a company. Participants read that "all 24 candidates are highly qualified and fulfill all the requirements the company is

looking for." Participants were instructed to take a careful look at the applicant pool and could only proceed to the next page after a minimum of 30 seconds. Then, participants saw eight applicants and were told to imagine that the company decided to hire these individuals. We manipulated the characteristics of the selected candidates. There were three between-subject conditions. In the attractiveness bias condition (n = 199), the selected candidates were balanced on gender (half male, half female) and race (half Black, half White), but only attractive candidates were selected. In the gender bias condition (n = 200), the selected candidates were balanced on attractiveness (half attractive, half unattractive) and race (half Black, half White), but only male candidates were selected. In the race bias condition (n = 200), the selected candidates were balanced on attractiveness (half attractive, half unattractive) and gender (half male, half female), but only White candidates were selected. Thus, participants viewed an attractiveness-biased, gender-biased, or race-biased selection of candidates. Participants were instructed to take a careful look at the selection and then "note down anything that stands out about the hiring process."

Response coding. Three research assistants who were blind to the research question, hypothesis, and experimental condition independently coded the responses of all participants. The primary dependent variable was whether participants mentioned attractiveness, gender, or race discrimination when describing the selection outcome. Coders indicated if the participant "thinks that recruiters used 'being attractive' ('being male'/'being white') as a criterion to select candidates from the original pool." Interrater reliability was high (Krippendorf's α = .74 for attractiveness, α = .97 for gender, α = .91 for race) and we coded participants' responses as mentioning a certain type of discrimination if at least two coders indicated this.

For exploratory purposes, the research assistants also coded if participants' responses included any reference to attractiveness, gender, or race (Krippendorf's α = .79 for attractiveness, α = .94 for gender, α = .93 for race; results are reported in the Supplementary Information).

Study 2

Participants. We recruited 199 first-year psychology students from a European university (M_{age} = 19.86 years, SD_{age} = 2.70; 83.84% female, 15.66% male, 0.51% non-binary) who completed the study in return for partial course credit. The sample size was determined by how many participants completed the study within seven weeks.

Stimuli and procedure. We used images of the same 24 targets varying in gender (male or female), race (Black or White), and attractiveness (low or high) as in Study 1. Participants saw the targets and were told to imagine that these individuals applied for a job with a company. Participants read that "all 24 candidates are highly qualified and fulfill all the requirements the company is looking for". Participants were instructed to take a careful look at the applicant pool and could only proceed to the next page after a minimum of 30 seconds. Then, participants saw 12 groups of eight applicants and were told to imagine that the company decided to hire these individuals. For each group, participants were asked to rate how fair they thought the selection process was on a scale ranging from 1 (*not fair at all*) to 7 (*extremely fair*).

There were four within-subject conditions with three trials per condition. As in Study 1, participants viewed attractiveness-biased, gender-biased, and race-biased selection. We also included an unbiased condition in which the selected candidates were balanced on attractiveness (half attractive, half unattractive), gender (half male, half female), and race (half Black, half White). Each bias type was shown three times for a total of 12 trials, and we used different combinations of targets on each trial. Trial order was fully randomized.

Finally, participants saw all 24 targets again and rated them on attractiveness on a scale ranging from 1 (*not attractive at all*) to 7 (*extremely attractive*).

Study 3

Study 3 was preregistered: https://aspredicted.org/KKG_JUA.

Participants. A power analysis in G*Power [46] indicated that, to detect a small-to-medium-sized effect ($d = 0.4$) with 80% power, 284 participants are required for a repeated-measures between-subjects ANOVA (3 measurements, 4 between-subjects conditions, $\alpha = 5\%$). Our main analyses of interest, *z*-tests comparing the attractiveness coefficient (reflecting the difference in fairness ratings in the attractiveness-biased vs. unbiased condition) with the race coefficient (reflecting the difference in fairness ratings in the race-biased vs. unbiased condition) and the gender coefficient (reflecting the difference in fairness ratings in the gender-biased vs. unbiased condition), might require more participants for adequate power. We thus targeted a sample size of 400 participants. We recruited 402 Prolific workers from the United States (M_{age} = 33.33 years, SD_{age} = 11.68; 47.26% female, 50.75% male, 1.50% non-binary; 70.90% White, 12.94% Black/African American, 1.49% American Indian/Alaska Native, 10.95% Asian, 0.50% Native Hawaiian/Pacific Islander, 9.20% Hispanic/Latinx, 0.50% other, 0% no response).

Stimuli and procedure. We implemented the same study design as in Study 2, with one exception: Type of bias was manipulated between subjects rather than within-subject. Participants were randomly assigned to one of four conditions that determined whether they saw attractiveness-biased (n = 101), gender-biased (n = 101), race-biased (n = 100), or unbiased (n = 100) candidate selections. After indicating their fairness ratings, participants in the attractiveness bias condition saw all 24 targets again and rated them on attractiveness on a scale ranging from 1 (*not attractive at all*) to 7 (*extremely attractive*).

Study 4

Study 4 was preregistered: https://aspredicted.org/QTC_YSG.

Participants. We implemented the same study design and analysis strategy as in Study 2, in which we found the predicted effect with a sample size of n = 199. As a conservative measure, we decided to target a sample size of 300 participants. We recruited 300 Prolific workers from the United States (M_{age} = 44.78 years, SD_{age} = 16.24; 50.33% female, 49.00% male, 0.67% non-binary; 66.33% White, 15.67% Black/African American, 2.33% American Indian/Alaska Native, 9.00% Asian, 0% Native Hawaiian/Pacific Islander, 9.00% Hispanic/Latinx, 2.67% other, 0.67% no response). Participants were selected to be representative of the general population (based on U.S. census data) along the dimensions of gender, age, and race.

Stimuli and Procedure. We used the same stimuli and a similar procedure compared to Study 2. Instead of describing a hiring scenario with different candidates being selected for a job, we used a legal sentencing scenario with different defendants being found guilty of a crime. We did not include a gender bias manipulation, and participants only considered male targets. Participants saw 20 male targets (half Black, half White; half attractive, half unattractive) and were told to imagine that these individuals were standing trial for the same type of alleged crime. Participants read that based on the available evidence, all defendants had approximately the same chance of being guilty of the crime. Participants were instructed to take a careful look at the defendant pool and could only proceed to the next page after a minimum of 30 seconds. Then, participants saw nine groups of eight applicants and were told to imagine that the judge decided to convict these individuals. For each group, participants were asked to rate how fair they think the sentencing process was on a scale ranging from 1 (*not fair at all*) to 7 (*extremely fair*).

There were three within-subject conditions with three trials per condition. In the attractiveness bias condition, the convicted individuals were balanced on race (half Black, half

White), but only unattractive defendants were found guilty. In the race bias condition, the convicted individuals were balanced on attractiveness (half attractive, half unattractive), but only Black candidates were selected. In the unbiased condition, the convicted individuals were balanced on attractiveness and race. Each bias type was shown three times for a total of nine trials, and we used different combinations of targets on each trial. Trial order was fully randomized.

After indicating their fairness ratings, participants in the attractiveness bias condition saw all 24 targets again and rated them on attractiveness on a scale ranging from 1 (*not attractive at all*) to 7 (*extremely attractive*).

Finally, we measured two variables for exploratory purposes. Participants completed a measure of social dominance orientation, i.e., the eight-item SDO7 Scale (e.g., "an ideal society requires some groups to be on top and others to be on the bottom")[38] and an adapted version [47] of the nine-item Everyday Discrimination Scale (e.g., "some people are treated with less courtesy than other people are") [39]. The adapted version measures the perceived prevalence of everyday discrimination, rather than personal experiences with everyday discrimination.

Study 5

Study 5 was preregistered: https://aspredicted.org/QSG_AJT.

Participants. A power analysis in G*Power indicated that, to detect a small-to-medium-sized interaction effect ($f = 0.175$) with 80% power and $\alpha = 5\%$, 318 participants are required. However, this estimate is based on implausible assumptions and others have suggested, as a rule of thumb, that the sample size needed to detect a 50% attenuation of an effect is approximately 14 times the sample size needed to detect the main effect [48]. In an earlier study, we found an effect size of $d = 1.3$ for the difference between attractiveness and race discrimination, which requires $n = 11$ per condition (a total of $n = 33$ across the three bias conditions) to detect the effect with 80% power and $\alpha = 5\%$. Thus, the 14x-rule would suggest a required sample size of $n = 462$ to detect an interaction effect with 50% attenuation. As a conservative measure, we decided to recruit 720 Prolific workers from the United States (M_{age} = 32.32 years, SD_{age} = 11.21; 48.19% female, 50.00% male, 1.81% non-binary; 71.81% White, 10.56% Black/African American, 1.67% American Indian/Alaska Native, 12.92% Asian, 0.69% Native Hawaiian/Pacific Islander, 7.78% Hispanic/Latinx, 0.97% other, 0% no response).

Stimuli and procedure. We used the same stimuli and implemented a similar study design as in Study 3. Participants saw a pool of 24 individuals that applied for a job and rated the fairness of different candidate selections on a scale ranging from 1 (*not fair at all*) to 7 (*extremely fair*). There were three between-subjects conditions that determined which type of selection bias participants were exposed to when viewing the group of selected candidates. In the attractiveness bias condition (n = 239), eight attractive candidates were selected. In the gender bias condition (n = 240), eight male candidates were selected. In the race bias condition (n = 241), eight White candidates were selected.

Additionally, participants were randomly assigned to one of two between-subject conditions. In the *control condition* (n = 363), participants underwent the same procedure as in Study 2. That is, participants saw the group of selected candidate selections and rated the fairness of the selection procedure with no additional information provided. In the *explicit discrimination condition* (n = 357), participants also saw a statement that made it explicit that the recruiter decided whom to hire based on the attractiveness, gender, or race of candidates (depending on the condition). For example, in the explicit attractiveness discrimination condition, participants read that "the recruiter decided to hire the most attractive candidates and reject the least attractive candidates".

Study 6

Study 6 was preregistered: https://aspredicted.org/9FB_3NJ.

Participants. A power analysis in G*Power indicated that, to detect a small-to-medium-sized interaction effect (d = 0.25) with 80% power in a 2 × 2 between-subjects design (with α = 5%), 505 participants are required. We recruited a sample of 505 Prolific workers from the United States (M_{age} = 29.15 years, SD_{age} = 9.95; 49.90% female, 48.32% male, 1.78% non-binary; 69.11% White, 13.47% Black/African American, 0.99% American Indian/Alaska Native, 7.33% Asian, 0.20% Native Hawaiian/Pacific Islander, 13.47% Hispanic/Latinx, 0.79% other, 0% no response). We used Prolific to target a sample that was representative with respect to gender, age, and race (based on U.S. census data).

Stimuli and procedure. We used a similar hiring context as in previous studies. In the current study, participants saw two highly qualified male candidates that are considered for a job. Below the images, participants also saw some additional background information on the two candidates, which was held constant across conditions (gender, age category, place of birth,

hobby, and college). Participants were asked to look at the profile of the two candidates. Then, they saw which candidate was selected for the job and we asked them to rate the extent to which different factors could have influenced the recruiter's decision. Of primary interest was the perceived role of candidates' attractiveness, but we also included six filler items that pertained to candidates' race, gender, age, place of birth, hobby, and college.

We implemented a 2 (attractiveness of rejected candidate: unattractive vs. attractive) × 2 (race of rejected candidate: White vs. Black) between-subjects design. For all participants, the selected candidate was White and attractive. We manipulated the attractiveness (unattractive: n = 253, attractive: n = 252) and the race (White: n = 253, Black: n = 252) of the candidate that was not selected. This allowed us to test if people are more likely to perceive attractiveness discrimination when a hiring decision *cannot* be explained by racial discrimination (i.e., when both the selected and rejected candidates are White) compared to when the decision *can* be explained by racial discrimination (i.e., when the selected candidate is White and the rejected candidate is Black).

Study 7

Study 7 was preregistered: https://aspredicted.org/9FJ_C5V.

Participants. A power analysis in G*Power indicated that, to detect a small-to-medium-sized difference in fairness ratings between the control condition and the intervention condition (d = 0.25) with 80% power (and α = 5%), 253 participants are required per condition. We, therefore, aimed to collect a total sample of 759 participants who satisfy our inclusion criteria. After excluding one participant who failed an attention check (as preregistered), a final sample of 761 Prolific workers from the United States remained (M_{age} = 33.34 years, SD_{age} = 12.82; 48.23% female, 49.54% male, 1.97% non-binary; 72.54% White, 8.80% Black/African American, 1.31% American Indian/Alaska Native, 11.70% Asian, 0.53% Native Hawaiian/Pacific Islander, 11.04% Hispanic/Latinx, 1.18% other, 0.39% no response). We used Prolific to target a sample that was gender-balanced.

Stimuli and procedure. We used the same stimuli and implemented a similar study design as in Study 2. Participants saw a pool of 24 individuals that applied for a job and rated the fairness of different candidate selections on a scale ranging from 1 (*not fair at all*) to 7 (*extremely fair*). There were four within-subject conditions determining which type of selection bias participants were exposed to when viewing the group of selected candidates. In the

attractiveness bias condition, eight attractive candidates were selected. In the gender bias condition, eight male candidates were selected. In the race bias condition, eight White candidates were selected. In the unbiased condition, selected candidates were balanced on attractiveness, gender, and race.

Participants were also randomly assigned to one of three between-subjects conditions (two intervention conditions and one control condition; all materials can be found here: https://osf.io/v2j3p/). In the *awareness condition* (n = 249), participants read an article about the prevalence and negative consequences of attractiveness discrimination (titled "The greatest privilege we never talk about: beauty"). The article was presented at the beginning of the study before participants read instructions for the main task. We used excerpts from real blog posts on appearance-based discrimination and the layout and design of the article were similar to popular science blogs. In the *debiasing condition* (n = 257), we explicitly instructed participants to pay attention to attractiveness discrimination. Participants first read the instructions for the main task and viewed the candidate pool. They were also told that hiring decisions often favor attractive people, that this type of discrimination often goes unnoticed, and that they should pay special attention to candidates' attractiveness. Participants then read the same article on attractiveness discrimination. When participants viewed the different groups of selected candidates and indicated their fairness ratings, we again instructed them to check for potential instances of attractiveness discrimination (next to other types of discrimination). In the *control condition* (n = 255), participants read an article of similar length about an unrelated topic ("The horseshoe crab"). Participants in all conditions completed an attention check question that asked them about the article's content immediately after it, with four answer options (attractiveness discrimination, the horseshoe crab, the Supreme Court, the Scottish Highlands).

Finally, participants saw all 24 targets again and rated them on attractiveness on a scale ranging from 1 (*not attractive at all*) to 7 (*extremely attractive*).

References

1. Nelson, T. D. *Handbook of Prejudice, Stereotyping, and Discrimination* 2nd edn (Psychology Press, 2016).
2. Mays, V. M., Cochran, S. D. & Barnes, N. W. Race, Race-Based Discrimination, and Health Outcomes Among African Americans. *Annu. Rev. Psychol.* **58**, 201-225, (2006).
3. Hebl, M., Cheng, S. K. & Ng, L. C. Modern Discrimination in Organizations. *Annu. Rev. Organ. Psychol. Organ. Behav.* **7**, 257-282, (2020).
4. Hinton, E. & Cook, D. The Mass Criminalization of Black Americans: A Historical Overview. *Annu. Rev. Criminol.* **4**, 261-286, (2021).
5. Bonezzi, A. & Ostinelli, M. Can algorithms legitimize discrimination? *J. Exp. Psychol. Appl.* **27**, 447-459, (2021).
6. Tomova Shakur, T. K. & Phillips, L. T. What counts as discrimination? How principles of merit shape fairness of demographic decisions. *J. Pers. Soc. Psychol.* **123**, 957-982, (2022).
7. Phillips, L. T. & Jun, S. Why benefiting from discrimination is less recognized as discrimination. *J. Pers. Soc. Psychol.* **122**, 825-852, (2022).
8. Kleinberg, J., Ludwig, J., Mullainathan, S. & Sunstein, C. R. Algorithms as discrimination detectors. *Proc. Natl Acad. Sci. USA* **117**, 30096-30100, (2020).
9. Barreto, M. & Ellemers, N. Detecting and Experiencing Prejudice: New Answers to Old Questions. *Adv. Exp. Soc. Psychol.* **52**, 139-219, (2015).
10. Hamermesh, D. S. *Beauty Pays: Why Attractive People Are More Successful* (Princeton University Press, 2011).
11. Dion, K., Berscheid, E. & Walster, E. What is beautiful is good. *J. Pers. Soc. Psychol.* **24**, 285-290, (1972).
12. Phillips, L. T., Jun, S. & Shakeri, A. Barriers and Boosts: Using Inequity Frames Theory to Expand Understanding of Mechanisms of Racial and Gender Inequity. *Acad. Manag. Ann.* **16**, 547-587, (2022).
13. Klebl, C., Luo, Y. & Bastian, B. Beyond Aesthetic Judgment: Beauty Increases Moral Standing Through Perceptions of Purity. *Pers. Soc. Psychol. Bull* **48**, 954-967, (2021).
14. Langlois, J. H. *et al.* Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* **126**, 390-423, (2000).
15. Talamas, S. N., Mavor, K. I. & Perrett, D. I. Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance. *PLoS ONE* **11**, e0148284, (2016).
16. Feingold, A. Good-looking people are not what we think. *Psychol. Bull.* **111**, 304-341, (1992).
17. Mobius, M. M. & Rosenblat, T. S. Why Beauty Matters. *Amer. Econ. Rev.* **96**, 222-235, (2006).
18. Hamermesh, D. S. & Biddle, J. E. Beauty and the Labor Market. *Amer. Econ. Rev.* **84**, 1174-1194, (1994).
19. Monk, E. P., Esposito, M. H. & Lee, H. Beholding Inequality: Race, Gender, and Returns to Physical Attractiveness in the United States. *Am. J. Sociol.* **127**, 194-241, (2021).
20. Nault, K. A., Pitesa, M. & Thau, S. The Attractiveness Advantage At Work: A Cross-Disciplinary Integrative Review. *Acad. Manag. Ann.* **14**, 1103-1139, (2020).
21. Graham, J. R., Harvey, C. R. & Puri, M. A Corporate Beauty Contest. *Manag. Sci.* **63**, 3044-3056, (2017).

22. Deng, W., Li, D. & Zhou, D. Beauty and job accessibility: new evidence from a field experiment. *J. Popul. Econ.* **33**, 1303-1341, (2020).
23. López Bóo, F., Rossi, M. A. & Urzúa, S. S. The labor market return to an attractive face: Evidence from a field experiment. *Econ. Lett.* **118**, 170-172, (2013).
24. Berggren, N., Jordahl, H. & Poutvaara, P. The looks of a winner: Beauty and electoral success. *J. Public Econ.* **94**, 8-15, (2010).
25. Mazzella, R. & Feingold, A. The Effects of Physical Attractiveness, Race, Socioeconomic Status, and Gender of Defendants and Victims on Judgments of Mock Jurors: A Meta-Analysis1. *J. Appl. Soc. Psychol.* **24**, 1315-1338, (1994).
26. Felton, J., Koper, P. T., Mitchell, J. & Stinson, M. Attractiveness, easiness and other issues: student evaluations of professors on Ratemyprofessors.com. *Assess. Eval. High. Educ.* **33**, 45-61, (2008).
27. Maestripieri, D., Henry, A. & Nickels, N. Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behav. Brain Sci.* **40**, e19, (2017).
28. Brooks, D. Why Is It OK to Be Mean to the Ugly? *New York Times (*24 June 2021).
29. Major, B. & Dover, T. L. in *Handbook of prejudice, stereotyping, and discrimination, 2nd ed.* 213-239 (Psychology Press, 2016).
30. Marti, M. W., Bobier, D. M. & Baron, R. S. Right before Our Eyes: The Failure to Recognize Non-Prototypical Forms of Prejudice. *Group Processes & Intergroup Relations* **3**, 403-418, (2000).
31. Inman, M. L. & Baron, R. S. Influence of prototypes on perceptions of prejudice. *J. Pers. Soc. Psychol.* **70**, 727-739, (1996).
32. Baron, R. S., Burgess, M. L. & Chuan Feng, K. Detecting and Labeling Prejudice: Do Female Perpetrators Go Undetected? *Pers. Soc. Psychol. Bull* **17**, 115-123, (1991).
33. Simon, S., Kinias, Z., O'Brien, L. T., Major, B. & Bivolaru, E. Prototypes of Discrimination: How Status Asymmetry and Stereotype Asymmetry Affect Judgments of Racial Discrimination. *Basic Appl. Soc. Psychol.* **35**, 525-533, (2013).
34. O'Brien, L. T., Kinias, Z. & Major, B. How status and stereotypes impact attributions to discrimination: The stereotype-asymmetry hypothesis. *J. Exp. Soc. Psychol.* **44**, 405-412, (2008).
35. Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proc. Natl Acad. Sci. USA* **105**, 11087-11092, (2008).
36. Wagenmakers, E.-J. A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* **14**, 779-804, (2007).
37. O'Brien, L. T., Major, B. & Simon, S. Why did you choose that person over me? Ingroup rejection and attributions to discrimination. *J. Exp. Soc. Psychol.* **48**, 1225-1233, (2012).
38. Ho, A. K. *et al.* The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *J. Pers. Soc. Psychol.* **109**, 1003-1028, (2015).
39. Williams, D. R., Yan, Y., Jackson, J. S. & Anderson, N. B. Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. *J. Health Psychol.* **2**, 335-351, (1997).
40. Inman, M. L., Huerta, J. & Oh, S. Perceiving Discrimination: The Role of Prototypes and Norm Violation. *Soc. Cogn.* **16**, 418-450, (1998).

41. Reynolds, T. *et al.* Man up and take it: Gender bias in moral typecasting. *Organ. Behav. Hum. Decis. Process.* **161**, 120-141, (2020).
42. Crosby, F., Clayton, S., Alksnis, O. & Hemker, K. Cognitive biases in the perception of discrimination: The importance of format. *Sex Roles* **14**, 637-646, (1986).
43. Axt, J. R., Casola, G. & Nosek, B. A. Reducing Social Judgment Biases May Require Identifying the Potential Source of Bias. *Pers. Soc. Psychol. Bull* **45**, 1232-1251, (2018).
44. Ma, D. S., Correll, J. & Wittenbrink, B. The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* **47**, 1122-1135, (2015).
45. Hehman, E., Xie, S. Y., Ofosu, E. K. & Nespoli, G. A. Assessing the Point at Which Averages Are Stable: A Tool Illustrated in the Context of Person Perception. *PsyArXiv*, (2018).
46. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175-191, (2007).
47. Kartolo, A. B. & Kwantes, C. T. Organizational culture, perceived societal and organizational discrimination. *Equal. Divers. Incl.* **38**, 602-618, (2019).
48. Blake, K. R. & Gangestad, S. On Attenuated Interactions, Measurement Error, and Statistical Power: Guidelines for Social and Personality Psychologists. *Pers. Soc. Psychol. Bull* **46**, 1702-1711, (2020).