Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes

People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance

--Manuscript Draft--

Manuscript Number:	PSP-I-2019-1536R3
Full Title:	People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance
Abstract:	Recognizing others' humanity is fundamental to how people think about and treat each other. People often ascribe greater humanness to groups that they socially value, but do they also systematically ascribe social value to different individuals? Here, we tested whether people (de)humanize individuals based on social traits inferred from their facial appearance, focusing on attractiveness and intelligence. Across five studies, less attractive and less intelligent-looking individuals seemed less human, but this varied by target gender: Attractiveness better predicted humanness attributions to women whereas perceived intelligence better predicted humanness attributions to men (Study 1). This difference seems to stem from gender stereotypes (preregistered Studies 2 and 3) and even extends to attributions of children's humanness (preregistered Study 4). Moreover, this gender difference leads to biases in moral treatment that confer more value to the lives of attractive women and intelligent-looking men (preregistered Study 5). These data help to explain how interpersonal judgments of individuals interact with intergroup biases to promote gender-based discrimination, providing greater nuance to the mechanisms and outcomes of dehumanization.
Article Type:	Article
Keywords:	attractiveness; dehumanization; face; gender; intelligence
Corresponding Author:	Ravin Alaei, PhD University of Toronto CANADA
Corresponding Author E-Mail:	ravin.alaei@mail.utoronto.ca
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Toronto
Other Authors:	Jason C. Deska, PhD
	Kurt Hugenberg, PhD
	Nicholas O. Rule, PhD
Author Comments:	
Corresponding Author's Secondary Institution:	
First Author:	Ravin Alaei, PhD
Order of Authors Secondary Information:	
Manuscript Region of Origin:	CANADA
Suggested Reviewers:	Eric Hehman, PhD Assistant Professor, McGill University eric.hehman@mcgill.ca Expert on how individuals perceive and evaluate one another across group boundaries (e.g., race, gender)
	Nour Kteily, PhD Associate Professor, Northwestern University Kellogg School of Management n-kteily@kellogg.northwestern.edu Expert on dehumanization and its role in intergroup and interpersonal relations

	Judith Hall, PhD Professor, Northeastern University j.hall@northeastern.edu Expert on the relations between gender, individual differences, and person perception
Opposed Reviewers:	
Order of Authors:	Ravin Alaei, PhD
	Jason C. Deska, PhD
	Kurt Hugenberg, PhD
	Nicholas O. Rule, PhD
Manuscript Classifications:	Group Processes; Intergroup Relations; Interpersonal Perception; Interpersonal Processes; Prejudice; Social Cognition; Social Identity; Stereotyping

Dr. Kteily's Comments:

1. **A critical feature of almost all papers in JPSP is that they provide not only evidence for the existence of a new phenomenon, but an explanation for it. All of the reviewers and I agree that there is something reasonably consistent going on in terms of attractiveness and intelligence predicting attributions of humanity (and your latest analyses are compelling in suggesting this isn't just due to likability). On the question of whether intelligence is a stronger predictor for attributions of men's humanity and attractiveness a stronger predictor for attributions of women's humanity, I would say that the preponderance of the evidence you provide across studies provides pretty solid support [although you do not make much of the finding in Study 3 that, presumably, attractiveness (standardized B = .68) matters more for attributions of boys' humanity than intelligence does (standardized B = .29)].**

   On the question of the reason for these patterns, I remain less convinced. Your introduction argues (quite compellingly) that the reason why intelligence would matter more for men than for women and attractiveness more for women than for men is because society places more value on men's intelligence and on women's attractiveness. The sole study that can speak to this as an *explanation* (beyond documenting that there *is* a difference in the link between attractiveness/intelligence and humanity across target gender) is Study 2, which assesses stereotype endorsement. You conclude (p. 24) that "Triangulating across the preregistered and exploratory analyses, these results indicate that gender stereotypes help account for why attractiveness better predicted humanness judgments of women and why perceived intelligence better predicted humanness judgments of men." But I struggled to come to this conclusion as I assessed your results. Part of this may have to do with what I found to be a confusing presentation of the exploratory analyses (see more below), but part of it might, frankly, simply have to do with the lack of empirical support for the assertions.

   To try to be as clear as possible, I will walk through the pieces of the argument as I encountered them (at the risk of being too detailed).

   I agree that Table 4 replicates the basic pattern from Study 1 that attractiveness matters more for women and intelligence matters for men. I also agree with the conclusion (p. 21) that for women, people who subscribed to the stereotype equating beauty and success for women dehumanized them more.

   But I found the evidence from the target*participant predictors part of Table 5 to be inconsistent with your assertions about specificity. Yes, it appears to be the case that "women's attractiveness more strongly influenced judgments of their humanness among participants who prioritized attractiveness in women compared to participants who did not" (i.e., the attractiveness*attractiveness stereotype endorsement moderation is significant, B=.48, p <.001). But it's also that women's *intelligence* more strongly influenced judgments of their humanness among participants who did vs didn't prioritize attractiveness. That is, the perceived

intelligence*attractiveness moderation in table 5 is also significant, and appears to be of similar magnitude. Your paper does not comment on this aspect of the results, even though it seems to me to importantly reduce confidence in the specificity of the claim that the *reason* why attractiveness matters more in evaluations for women is because of greater endorsement of attractiveness stereotypes.

The same is true when turning to analyses for the men (Table 6), to an even greater extent. You find the predicted moderation (perceived intelligence*intelligence stereotype endorsement; B=1.00, r= .02). But you also find that perceptions of attractiveness matter more for those who endorse the intelligence stereotype (and the effect size for this perceived attractiveness*intelligence moderation, B=.80, r=.03, is just as large if not larger than the predicted one). Moreover, it's also the case that the attractiveness*attractiveness stereotype endorsement effect for men (B=.45, r=.02) is significant and seems potentially comparable in effect size to that same interaction among women (B=.48, r=.03). Your conclusions say too little about this complexity and lack of specificity. Of the potentially problematic interactions I mention (problematic for your case about stereotype endorsement as an explanation for target gender differences), the only one you mention is the attractiveness*attractiveness stereotype among men….

…This is particularly critical because as noted before, it is important for a paper published in JPSP to be able to provide evidence for the reasons behind a phenomenon it uncovers. I see this as the primary remaining barrier to publication (I am happy with the way you have addressed the others, despite some limitations). It may be that there is some way that you can make a more convincing case with the data that you already have from Study 2 despite what appear to me to be some theoretically-inconsistent patterns. One possible analysis that occurred to me that might help: in your section exploring analyses of participant gender effects in Study 2, you note differences like female raters using female targets' attractiveness (but not intelligence) more than male raters did when rating female targets' humanness. Can you show evidence, perhaps, that this is mediated by gender differences in endorsement of the attractiveness stereotype for women?

But I think you might also consider whether it could make sense to add some data that makes the case more straightforwardly, in which you could perhaps also build upon some of the previously noted limitations of your stereotype endorsement measure and build in pre-registered predictions that speak directly to those aspects of your exploratory analyses that you think are most compelling in providing evidence for your claims. I will leave it to you to determine the best means to achieve the important goal of making a clearer case for mechanism.

Thank you for your feedback. As you will see in the manuscript, we took the long road here by collecting new data and greatly appreciate your patience with us as we did so.

First, upon inspecting the results from Study 2 in totality, we believe that they generally support our previous theorizing (especially when considering the exploratory analyses, which we hope to have successfully clarified; please see below). Namely, valuing attractiveness resulted in

generally dehumanizing women (but not men), whereas valuing intelligence resulted in generally dehumanizing men (but not women). Furthermore, apart from the two interaction terms noted in your comment, the remaining interaction terms fit the hypothesis (including the attractiveness × attractiveness stereotype endorsement interaction effect in the male targets model; again, please see our clarification of the exploratory analyses). Thus, considering the whole of the seven preregistered effects and four exploratory effects, the bulk of the evidence does seem to support the conclusions. In other words, despite the imperfections introduced by the two questionable interaction terms that you described so well above, we feel confident that the conclusions we draw from the longer arc of the data are real: though the collection of results is a bit noisy, we feel affirmed that 85% of the effects testing the mechanism for these effects all point in the hypothesized direction, rendering it very unlikely that they are simply based on persistent Type I errors.

We nevertheless also agree that those two interaction terms are difficult to interpret, and thus recognize that having had to create a scale invites the possibility that we were not precisely measuring what we aimed to measure. If we understand your suggestion correctly regarding testing the mechanism via the participant gender data, we unfortunately could not find a way to model it (i.e., it requires a moderated mediation of a cross-classified model, an analysis that seems beyond any currently available software packages that we know).

We therefore conducted a new study with another measure (as mentioned) and meta-analyzed the results across Study 2 and the new Study 3 to examine which effects remain robust. For a full description, please see Study 3 in the revised manuscript; but, in brief, we believe that the results from Study 3 and the meta-analysis further show that gender-stereotyped-valuing of women's beauty underlies why their attractiveness more strongly influences their judged humanness compared to men's attractiveness. Admittedly, the results for male targets are more mixed, though the combined meta-analytic result continues to show that valuing intelligence only influenced ratings of men's humanness. We speculate that the remaining inconsistent results for the male targets stem from gender subtyping, which we explain in the Discussion section of Study 3 (and again in the General Discussion).

2. **You include exploratory analyses that you argue help to probe the question of whether gender stereotypes explain attractiveness judgments of women's humanness more than men's. I confess that despite reading this exploratory section several times, I found your description of the analyses confusing, and struggled to fully understand how they can help convince readers about your theoretical argument in light of the points noted above.**

   You state (pp. 22-23) that you "added target gender to the interaction terms that had indicated the importance of gender stereotypes, allowing us to test whether gender stereotypes influence one gender over the other for a particular trait among participants who highly valued that trait". For one, I couldn't tell specifically what models you are testing here. Are you suggesting that you tested the same models as in tables 5-6 but collapsing them into one large analysis by adding target gender as an interaction term to all of the variables in the tables? Are you instead suggesting that you looked at gender*participant predictors (i.e., gender*attractiveness stereotype endorsement; gender*intelligence stereotype endorsement), ignoring the

role of the targets' rated attractiveness and perceived intelligence? Or that you added gender to just the target*participant predictors? (i.e., for example, gender*attractiveness*attractiveness stereotype endorsement, and so on)?

I think including a specific table here that very clearly indicates what is being tested is important. Second, I think spelling out your logic as you set up the analysis with a specific example that highlights the relevance of this analysis to your theory would be clearer than "allowing us to test whether gender stereotypes influence one gender over the other for a particular trait among participants who highly valued that trait." (And I'd encourage you to do this and then give the results for this set of analyses before turning to 'absolute magnitude' to help avoid further confusion).

In describing the results for this 'relative' analysis, you state that the target gender significantly interacted with intelligence but not attractiveness, but I can't tell if you are referring to intelligence and attractiveness stereotype endorsement or target intelligence and attractiveness. If, as I suspect, it is indeed stereotype endorsement, it's also not clear to me how you are treating this across men and women. You say "the relative difference in humanness ratings between participants who highly valued attractiveness and those who did not was roughly equal regardless of whether they rated male or female targets", but this sounds like you are classifying participants on whether they, as people, valued attractiveness in general. This seems inconsistent with the stereotype endorsement measure, which is always with reference to whether you think attractiveness is important *for a given gender*.

Even taking the analyses for granted, your conclusion from the 'relative analyses' that the difference between participants who highly valued attractiveness and those who did not was roughly equal regardless of whether they rated male or female targets does not seem to support the conclusion that attractiveness matters more when judging women because of societal stereotypes about women being valued because of their looks more than men. Although here too I was thrown off, because when I turned to the 'absolute' analyses (which, again, I could not follow clearly for very similar reasons), I saw that you concluded that "among participants who highly valued attractiveness, women's attractiveness influenced their humanness judgments significantly more than men's". Unless I'm missing something, this seems at odds with the conclusion from the relative analyses.

In sum, I found the section on the exploratory analyses quite confusing, and did not think your paper was sufficiently clear in terms of clarifying how these analyses were run, and why these analyses support your theory about stereotype endorsement underlying the gender differences you posit despite some of the primary analyses (target*participant predictor interactions in tables 5 and 6) painting a muddier picture.

Reading this section again, we agree that it could have been written more clearly. As you suggest, we have now (i) added examples to clarify the reasoning behind the exploratory analyses, (ii) explicated the statistical models underlying the exploratory analyses, and (iii)

reordered the reporting of the exploratory analyses such that the relative magnitude results are reported first in their entirety, followed by the absolute magnitude results. We refrained from including a table to report the exploratory results because we worried that it would be unwieldy (with four three-way interaction terms for each target gender) and repetitive (considering that many of the effects are already presented in the other tables). Still, we hope that the revisions now illuminate the reason for the exploratory analyses; we would be happy to add the table if you believe it would help readers better understand the analyses:

> From pages 26-30 (Study 2 Results):
> Thus far, the results suggest that gender stereotype endorsement can help to explain why attractiveness and perceived intelligence differently influence judgments of men's and women's humanness. Yet, we also found that attractiveness influenced humanness ratings among participants who reported highly valuing men's attractiveness, $b = 3.57$, $SE = 0.51$, $t(353) = 7.05$, $p < .001$, $r_{\text{effect size}} = .35$, compared to participants who did not, $b = 0.82$, $SE = 0.35$, $t(380) = 2.37$, $p = .02$, $r_{\text{effect size}} = .12$. This may indicate that beliefs about the importance of attractiveness similarly influence humanness judgments for both men and women. In other words, gender stereotypes may not explain why attractiveness affects judgments of women's humanness more than men's.
>
> Two exploratory comparisons helped us to probe this question and further substantiate our conclusion regarding these gender differences. First, we created a model with targets' gender; group-mean-centered attractiveness, perceived intelligence, and likeability (using the consensus scores from Study 1); and interactions between target gender, target attractiveness/perceived intelligence, and participants' stereotype endorsement (i.e., target gender × target attractiveness × attractiveness stereotype endorsement, target gender × target attractiveness × intelligence stereotype endorsement, target gender × target perceived intelligence × attractiveness stereotype endorsement, target gender × target perceived intelligence × intelligence stereotype endorsement) to predict the humanness ratings in a cross-classified model. The three-way interaction terms (e.g., target gender × target perceived intelligence × intelligence stereotype endorsement) test whether the difference in influence between those who value a trait highly and those who do not value a trait is greater for ratings of one target gender over the other.
>
> To illustrate, imagine that (when rating female targets' humanness) people who highly value women's intelligence are influenced by the female targets' apparent intelligence twice as much as people who do not value women's intelligence. Further imagine that (when rating male targets' humanness) people who highly value men's intelligence are influenced by the male targets' apparent intelligence four times as much as people who do not value men's intelligence. In this hypothetical case, targets' perceived intelligence interacts with participants' endorsement of the intelligence stereotype for ratings of both target genders, indicating that the degree to which a participant values others' intelligence, whether male or female, influences how much they regard (un)intelligent-looking looking men and women as human. This may seem to indicate that the endorsement of the intelligence stereotype does not account for why male targets' humanness depends more on their apparent intelligence compared to female targets. Yet, it does actually account for why men's humanness depends more on their perceived intelligence than women's humanness does because the difference between people who highly value (men's/women's) intelligence and people who do not highly value intelligence is twice as large for the male versus female targets (i.e., four times as much for male targets versus two times as much for female targets). Thus, the *relative* magnitude of the importance of the trait differs between judgments of men and women.
>
> In assessing relative magnitude, the three-way interaction for attractiveness was not significant (i.e., target gender × target attractiveness × attractiveness stereotype

endorsement), $b = -0.03$, $SE = 0.15$, $t(42010) = -0.19$, $p = .85$, $r_{\text{effect size}} = .00$, whereas the three-way interaction for intelligence was significant (i.e., target gender × target perceived intelligence × intelligence stereotype endorsement), $b = 1.06$, $SE = 0.29$, $t(42010) = 3.69$, $p < .001$, $r_{\text{effect size}} = .02$. Thus, the relative difference in humanness ratings between participants who highly value attractiveness and participants who do not highly value attractiveness was similar (i.e., not significant) for their ratings of male and female targets. Conversely, the relative difference in humanness ratings between participants who highly value intelligence and participants who do not highly value intelligence was greater for male targets than for female targets.

These relative magnitude analyses help to further substantiate that intelligence stereotype endorsement accounts for why perceived intelligence influences ratings of male targets' humanness more than female targets' humanness. However, they do not explain why attractiveness influences ratings of female targets' humanness more than male targets' humanness. Because the analyses only compare relative differences between genders among participants who highly value a trait versus participants who do not highly value that trait, the possibility that participants who highly value a trait are more influenced by it in *absolute* terms for one gender versus another gender remains open. In other words, participants who highly value attractiveness likely value it in both men and women, but also likely more so in women. Conversely, participants who do not highly value attractiveness likely do not particularly value it in either men or women, but still more so in women. Analogous differences should apply to intelligence.

To further illustrate: imagine that the standardized regression coefficient for how much target attractiveness predicts ratings of male targets' humanness among people who highly value men's attractiveness is $\beta = .25$, whereas it is $\beta = .10$ for people who do not value men's attractiveness. Further imagine that the standardized regression coefficient for how much target attractiveness predicts ratings of female targets' humanness among people who highly value women's attractiveness is $\beta = .50$, whereas it is $\beta = .20$ among people who do not value women's attractiveness. In each case, endorsement of the attractiveness stereotype influences the degree to which men's and women's attractiveness affects ratings of their humanness (because the standardized regression coefficients are larger in magnitude among people who highly value attractiveness vs. people who do not for both genders; i.e., .25 vs. .10 for male targets, and .50 vs. .20 for female targets). Indeed, the relative magnitude is multiplicatively equal (.25 is 2.5 times larger than .10, and .50 is 2.5 times larger than .20). But the absolute magnitude for the female targets is double that for the male targets (i.e., .50 vs. .25, and .20 vs. .10). Thus, absolute differences in effect sizes can also signal differences in the impact of stereotype endorsement across genders. To better capture these absolute differences, we conducted a second series of exploratory analyses comparing the effect sizes that measure how strongly each trait predicts target gender differences among participants who highly value the trait and among participants who do not highly value the trait (i.e., the effect sizes ascertained from the simple slopes analyses of the target × participant interaction terms from the models outlined in Tables 5 and 6).

Among participants who highly value attractiveness, women's attractiveness ($r_{\text{effect size}} = .60$) influenced their humanness judgments significantly more than men's attractiveness did ($r_{\text{effect size}} = .35$), meta-analytic $Z = 4.62$, $p < .001$. Similarly, among participants who do not highly value attractiveness, women's attractiveness ($r_{\text{effect size}} = .46$) still influenced their humanness judgments more than men's attractiveness did ($r_{\text{effect size}} = .12$), meta-analytic $Z = 5.25$, $p < .001$.

Conversely, among participants who highly value intelligence, men's perceived intelligence ($r_{\text{effect size}} = .42$) influenced their humanness judgments significantly more than women's perceived intelligence did ($r_{\text{effect size}} = .23$), meta-analytic $Z = 3.06$, $p = .002$. Finally, among participants who do not highly value intelligence, men's perceived

> intelligence ($r_{\text{effect size}}$ = .19) did not influence humanness judgments more than women's perceived intelligence did ($r_{\text{effect size}}$ = .26), meta-analytic Z = 1.05, p = .29. Triangulating across the preregistered and exploratory analyses, these results indicate that gender stereotypes help to account for why attractiveness better predicted judgments of women's humanness and why perceived intelligence better predicted judgments of men's humanness.

3. **Finally, I noticed that there continue to be what appear to be statistical errors. On p. 15, you refer to a marginally significant association between intelligence and humanity referencing Table 2, but the association in Table 2 appears to be significant (p < .001). On p. 34, you refer to "attractiveness only marginally did (Table 8)", but Table 8 shows that the p-value for attractiveness is .02. I would urge you to comb through the manuscript to remove all statistical errors.**

We apologize for these errors. We have gone through the manuscript and ensured that we have accurately reported all analyses and results.

Dr. Nour Kteily
Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes

December 31, 2020

Dear Dr. Kteily,

My co-authors and I again thank you for giving us the opportunity to revise our manuscript and for a generous extension of the deadline. In response to your helpful feedback, we conducted a new study using another measure of gender stereotyping to further clarify the mechanism underlying our results. Furthermore, we ensured accuracy in reporting all statistical analyses in the manuscript and revised our description of the exploratory analyses in Study 2 to be more straightforward. We are thus pleased to resubmit our revised manuscript entitled "People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance" for further consideration at the *Journal of Personality and Social Psychology*.

As before, the data collection and treatment of participants both followed the guidelines prescribed by the American Psychological Association. Furthermore, none of us has any conflict of interest regarding these data, which have not been previously published, and we all agreed to this submission. It is my pleasure to continue serving as the corresponding author for this manuscript.

I hope that you will enjoy reading the revised copy of our work and that you will agree with our assessment that it presents a contribution of interest to the readership at the *JPSP: IRGP*.

Sincerely,

Ravin Alaei
Department of Psychology
University of Toronto
Sidney Smith Hall
Toronto, ON M5S 3G3
ravin.alaei@mail.utoronto.ca

Jason C. Deska
Department of Psychology
Ryerson University
350 Victoria Street
Toronto, ON M5B 2K3
jdeska@ryerson.ca

Kurt Hugenberg
Department of Psychological and Brain Sciences
Indiana University
Bloomington, IN 47405-7007
khugenb@iu.edu

Nicholas O. Rule
Department of Psychology
University of Toronto
Sidney Smith Hall
Toronto, ON M5S 3G3
rule@psych.utoronto.ca

People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance

Ravin Alaei,[a] Jason C. Deska,[a] Kurt Hugenberg,[b] and Nicholas O. Rule[a]

[a] University of Toronto

[b] Indiana University

Author Note

Ravin Alaei, Department of Psychology, University of Toronto; Jason C. Deska, Department of Psychology, University of Toronto (email: jason.deska@utoronto.ca; phone: 1-416-946-0411); Kurt Hugenberg, Department of Psychological and Brain Sciences, Indiana University (email: khugenb@iu.edu; phone: 1-812-855-2151); Nicholas O. Rule, Department of Psychology, University of Toronto (email: rule@psych.utoronto.ca; phone: 1-416-978-3948).

Correspondence concerning this article should be addressed to Ravin Alaei, 100 St. George Street, Department of Psychology, University of Toronto, Toronto, ON, M5S 3G3. Email: ravin.alaei@mail.utoronto.ca.

People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance

Submission Date: January 17, 2019

Re-submission Date: February 14, 2020

Re-submission Date: December 31, 2020

Abstract

Recognizing others' humanity is fundamental to how people think about and treat each other. People often ascribe greater humanness to groups that they socially value, but do they also systematically ascribe social value to different individuals? Here, we tested whether people (de)humanize individuals based on social traits inferred from their facial appearance, focusing on attractiveness and intelligence. Across five studies, less attractive and less intelligent-looking individuals seemed less human, but this varied by target gender: Attractiveness better predicted humanness attributions to women whereas perceived intelligence better predicted humanness attributions to men (Study 1). This difference seems to stem from gender stereotypes (preregistered Studies 2 and 3) and even extends to attributions of children's humanness (preregistered Study 4). Moreover, this gender difference leads to biases in moral treatment that confer more value to the lives of attractive women and intelligent-looking men (preregistered Study 5). These data help to explain how interpersonal judgments of individuals interact with intergroup biases to promote gender-based discrimination, providing greater nuance to the mechanisms and outcomes of dehumanization.

Keywords: attractiveness; dehumanization; face; gender; intelligence.

People Attribute Humanness to Men and Women Differently Based on Their Facial Appearance

Recognizing other people as fellow humans forms a critical foundation for a civil society. Acknowledging others' humanness validates their moral capacities and fundamental rights, whereas dehumanizing others promotes intergroup conflict, immoral treatment, discrimination, and aggression (Gray et al., 2007; Haslam & Loughnan, 2014; Opotow, 1990; Pereira et al., 2009; Viki et al., 2013). People more often recognize the humanity of those they value than those they do not (e.g., Cikara et al., 2010; Sherman & Haidt, 2011). In particular, individuals who belong to stigmatized or devalued groups often endure both overt and subtle dehumanization (Harris & Fiske, 2006; Kteily et al., 2015).

Whereas devalued groups often seem less than fully human (Haslam & Loughnan, 2014), research shows a surprising amount of intragroup variability in how much perceivers ascribe sophisticated human faculties to others, even based on facial cues. For example, people ascribe less sophisticated mental capacities to targets displaying expressions of rejection (e.g., disgusted expressions) versus acceptance (e.g., happy expressions), consequently seeing them as less human (Krumhuber et al., 2018; see also Powers et al., 2014). Similarly, targets with direct eye gaze garner more human mental attributions than targets with averted gaze (Khalid et al., 2016). Conceptually similar effects occur for variations in facial structure. Targets with more dominant faces (driven by a higher width-to-height ratio) seem less human than targets with more submissive faces (Deska et al., 2018). Ascriptions of sophisticated humanlike faculties thus vary across targets, and the targets' facial cues promote these differences.

Extending this logic, we investigated the novel question of whether inferences of socially valued traits can affect ascriptions of humanness—specifically, appearing facially attractive and intelligent. In other words, we wondered whether people who look attractive and intelligent may

seem more human. We therefore first investigated whether these particularly valued human traits trigger stronger attributions of humanness; and second, whether these inferences differ depending on the target's gender. Because society often values women for their beauty and men for their intelligence, we queried whether more attractive women and more intelligent-looking men would seem more fully human.

In addressing these questions, we first outline the existing research and theory on precursors to dehumanization before turning to a discussion of how society values women and men for beauty and intelligence, respectively. We then report five studies in which we investigated the disparate dehumanization of men and women via traits perceived from their faces.

Precursors of Dehumanization

Deciding whether another person is human sounds trivial. Whether operationalized in terms of humanlike minds (e.g., Rai et al., 2017; Waytz et al., 2010) or as something more (e.g., Fincher et al., 2018), it seems intuitive that others are animate and have sophisticated human capacities, such as an intellect and the capacity for love. Indeed, simply perceiving humanlike facial features and configurations can spontaneously trigger the categorization "human" (Deska et al., 2017; Looser & Wheatley, 2010). But this seemingly automatic response actually masks a constructive process in which people integrate their perceptions of others with their own motivations (e.g., the desire to predict their behavior, connect with them, or to bolster their ingroups; see Deska & Hugenberg, 2017, and Haslam & Loughnan, 2014, for reviews).

Such selfish needs can often determine whether one ascribes humanlike mental and emotional faculties to others. Put simply, people appear to ascribe and withhold humanness when it suits them. For instance, people often dehumanize groups that possess low social value or

stature (e.g., the homeless and drug addicts; Harris & Fiske, 2006, 2007; Sherman & Haidt, 2011). People higher in social power and status similarly deny the humanity of devalued individuals below them (Capozza et al., 2012; Gwinn et al., 2013). Likewise, people often deny outgroup members' capacity to experience complex human emotions, such as love and ennui (especially for stigmatized outgroups; e.g., refugees, minorities, and immigrants; see Leyens et al., 2007, for a review). Reciprocally, people often see others as having sophisticated, humanlike minds when doing so fulfills their needs (e.g., social belongingness; Powers et al., 2014). Data from multiple sources thus suggest that perceiving others' humanness may arise from motivated perception, especially intergroup motives. People routinely dehumanize groups (and individuals) that they devalue and ascribe human faculties to individuals or groups depending on their goals.

Much of the extant research examining dehumanization has adopted an intergroup perspective: Residents dehumanize the homeless, the powerful dehumanize the powerless, the majority group dehumanizes minority groups, and citizens dehumanize immigrants and refugees. Although dehumanization has historically occurred and continues to occur in intergroup contexts (facilitating group-based conflict; Kteily & Bruneau, 2017), emerging research indicates that variability among group members' characteristics cue *individuals'* relative humanity too (Deska & Hugenberg, 2017). This dovetails with accumulating evidence showing that facial appearance can affect how human one seems (Deska & Hugenberg, 2017). For instance, people ascribe more humanlike minds to faces with more humanlike features (e.g., doll-like vs. human-like eyes), to faces with more humanlike structures (e.g., narrower and longer vs. wider and shorter), and to faces that expedite configural processing (i.e., presented upright vs. inverted; Deska et al., 2018; Hugenberg et al., 2016; Looser & Wheatley, 2010). Here, we extended this work on social cues

to humanness by examining whether judgments of individuals' humanity depends on the perception of two key traits: attractiveness and intelligence.

Trait- and Gender-Specific Cues to Humanness

Person judgment often begins with the very first impressions one makes from others' faces (Hehman et al., 2017). These initial impressions subsequently guide interactions, scaffolding how they unfold and the outcomes they yield (Gunaydin et al., 2017). An individual's physical or facial appearance can therefore play an outsized role in how others view and treat them (see Zebrowitz, 2017, for review). Whether accurate or not, people infer others' traits from facial appearance with both great speed and consensus (e.g., Rule et al., 2013; Todorov et al., 2009). These trait inferences can influence important real-world decisions, such as whether to hire a prospective employee, to warrant the use of police force, or even to execute a convicted murderer (e.g., Rule et al., 2016; Wilson et al., 2017; Wilson & Rule, 2015).

As noted above, these face-based trait inferences can also elicit different judgments of individuals' humanness (Deska & Hugenberg, 2017). Attractiveness and intelligence are particularly important in face judgments (e.g., Zebrowitz et al., 2002). Thus, we were particularly interested in whether facial attractiveness and intelligence influence judgments of others' humanness. Yet, because gender stereotypes prize women for their beauty and men for their intellect and agency (e.g., Smiler & Epstein, 2010), the dehumanization of unattractive and unintelligent-looking people may differ depending on the target's gender. We outline below how face-based perceptions of attractiveness and intelligence each drive judgments of humanness in potentially different directions for male and female targets.

Gender's Role in Ascribing Humanness from Facial Attractiveness. Facial attractiveness automatically orients attention, enjoys broad cross-cultural consensus, and has

potent downstream consequences for trait impressions (Eagly et al., 1991; Olson & Marshuetz, 2005; Sui & Liu, 2009). Perhaps more relevant to the present work, some forms of physical attractiveness have already stimulated hypotheses about how perceivers ascribe humanness to targets.

Sherman and Haidt (2011) theorized that physically *cute* targets elicit greater inferences of a humanlike mind because they motivate social engagement, which itself triggers ascriptions of mind. Although this hypothesis lacks direct experimental evidence, it does have indirect support. For example, many easily anthropomorphized objects have facial cues that exaggerate cuteness, such as large eyes in cartoon characters and children's dolls.

Indeed, attractive individuals may appear to carry more social value than unattractive individuals for a variety of reasons. More attractive targets signal greater mate value, mental stability, and social responsiveness (Farina et al., 1977; Fink & Penton-Voak, 2002). Multiple theories about the characteristics of *humanness* indicate that cognitive capacity, civility, emotional sensitivity, self-control, vitality, and warmth form essential components of perceiving someone as human—all characteristics ascribed to physically attractive individuals (Dion et al., 1972; Haslam, 2006; Waytz et al., 2010; Wilson et al., 2018).

Conversely, people often respond to unattractive individuals in dehumanizing ways. For instance, people attend to attractive faces and ignore unattractive faces (Rubenstein et al., 1999; Sui & Liu, 2009). Moreover, they feel disgusted by unattractive faces (much as they do by various stigmatized groups), leading them to cognitively and emotionally disengage from them (Krendl et al., 2006; Principe & Langlois, 2009; Schein & Langlois, 2015). Thus, people behaviorally, cognitively, and emotionally eschew unattractive individuals, potentiating dehumanization.

The effects of physical attractiveness on judgment are often stronger for women than for men, however. Appearing attractive can facilitate women's educational, professional, and social lives more than men's (Feingold, 1990; Fredrickson & Roberts, 1997; Striegel-Moore et al., 1986; Unger, 1979). This stems partly from the perception of unattractive women as especially disgusting (Tiggeman & Lewis, 2004). These associations may explain why women and girls often internalize the importance of attractiveness such that it affects their self-concept more than men's, amplifying self-objectification and generating substantial pressure to focus on their appearance (Gillen & Lefkowitz, 2009; Lerner et al., 1976; Tiggemann & Lynch, 2001). Debates about the reasons for this greater emphasis on women's attractiveness notwithstanding, attractiveness does appear to carry more value for women than for men (e.g., Buss, 1989; Gillen & Lefkowitz, 2009; Jackson, 1992).

Insofar as social value motivates ascriptions of humanness, and insofar as women's perceived social value greatly depends on their attractiveness, facial attractiveness may affect ascriptions of humanness especially strongly for female targets (Cikara et al., 2010; Epley et al., 2007; Sherman & Haidt, 2011). We therefore hypothesized that attractiveness would influence humanness judgments for women more than for men, first exploring this hypothesis in Study 1 and then conducting preregistered direct and conceptual replications in Studies 2-5.

Gender's Role in Ascribing Humanness from Facial Intelligence. Similar to attractiveness, people form intelligence impressions from faces quickly and with great consensus in ways that generate potent social consequences (e.g., children's academic performance, politicians' electoral success; Antonakis & Dalgas, 2009; Rosenthal & Jacobson, 1968). Looking intelligent thus has real social value: Unintelligent individuals seem less capable of achievement,

seem less worthy as mates, and generate greater contempt and disgust that leads people to disengage from them (Hutcherson & Gross, 2011; Li et al., 2002; Miceli & Castelfranchi, 2017).

Given the large range of outcomes indicating that people consider unintelligent individuals incapable and unworthy, we hypothesized that facial intelligence may also signal humanness. Put simply, individuals whose faces appear more intelligent may seem more fully human. Indeed, multiple psychological theories stress the link from the capacity for thinking and reasoning to ascriptions of humanness (e.g., Gray et al., 2007; Haslam, 2006). For example, Haslam's (2006) Dual Model of Dehumanization argues that humans' capacity for rationality and logic uniquely separates them from animals. Theories of mind perception also identify *agency*—the capacity for thought, planning, and self-regulation—as a critical dimension of mind (Gray et al., 2007). Furthermore, early work on "infrahumanization theory" identified intelligence as the most commonly nominated human characteristic (Leyens et al., 2001). Past literature therefore converges on the conclusion that intelligence may form a core component of humanness ascriptions. Based on these insights, we predict that people will ascribe more humanness to targets with more (vs. less) intelligent-looking faces.

Whereas previous research supports the hypothesis that intelligence undergirds judgments of humanness (see Gray et al., 2007; Haslam, 2006; Leyens et al., 2001), this too may vary by target gender. Intelligence features prominently in stereotypes distinguishing men from women (Smiler & Epstein, 2010). This distinction influences domestic and professional decisions throughout the lifecycle, accentuating the need for men to display intelligence and competence in order to establish and maintain their value as mates (e.g., Broverman et al., 1972; Prokosch et al., 2009; Smith & Wilhelm, 2004). In turn, men typically judge their intelligence as higher than women's (Furnham & Rawles, 1995). Reciprocally, when imagining an "intelligent"

individual, people describe a man and masculine traits (Raty & Snellman, 1992). Expectations linking intelligence with masculinity extend to many professional fields, in which members believe that only men possess the "natural brilliance" required to succeed (Leslie et al., 2015; Meyer et al., 2015). Men express acute awareness of the expectations for their intelligence, thus showing more susceptibility than women to impostor syndrome (the feeling that others will soon discover that one actually lacks the intelligence needed to attain one's existing status; Badawy et al., 2018). In other words, people expect men to demonstrate intelligence and achievement more than women.

Because society values men's intelligence more than women's intelligence, we hypothesized that this different emphasis may lead to distinct patterns of dehumanization based on gender. Specifically, although intelligence forms an essential component of humanness (e.g., Haslam, 2006), we expected that it would influence perceptions of men's humanness more than women's. We tested this hypothesis in Study 1, followed by preregistered replications in Studies 2-5.

Current Work

Here, we investigate how interindividual (rather than intergroup) variations in facial attractiveness and intelligence influence ascriptions of humanness. We hypothesized a priori that more attractive and more intelligent-looking targets would seem more fully human (Studies 1-4), and that their lives would carry more value (Study 5). We also explored whether the importance of these facial cues would vary by gender in Study 1 (i.e., whether there would be a stronger association between attractiveness and humanness among female targets in contrast to a stronger association between intelligence and humanness among male targets). Upon confirming these hypotheses, we subsequently replicated the results in four confirmatory studies to examine

potential mechanisms (Studies 2 and 3), to expand their scope to include judgments of children (Study 4), and to investigate downstream consequences that might ensue (Study 5).

Study 1

Studies have shown that a variety of labile (e.g., expression, eye gaze) and fixed (e.g., width-to-height ratio) facial cues can influence the extent to which targets are ascribed fully humanlike faculties (see Deska & Hugenberg, 2017, for review). We extended this logic, testing whether variation in facial cues of attractiveness and intelligence influences ascriptions of individuals' humanness.

In Study 1, we investigated whether one's facial attractiveness and facial intelligence would relate to ascriptions of how human the person seems. Because humanness ascriptions involve recognizing others' capacity to think and reason (e.g., Gray et al., 2007; Haslam, 2006), and facial attractiveness governs ascriptions of myriad positive social qualities (e.g., Dion et al., 1972), interindividual variation in perceptions of intelligence and attractiveness seemed to be especially important potential sources for trait-based ascriptions of humanness. We therefore hypothesized that perceptions of targets' attractiveness and intelligence would correlate with how human they seem.

Noting that attractiveness and intelligence possess different value for men and women, we also tested whether target gender modulates the association between perceptions of attractiveness, intelligence, and humanness. The premium placed on women's attractiveness and men's intelligence might lead to stronger attractiveness-humanness links for women than for men, and stronger intelligence-humanness links for men than for women. We therefore conducted these analyses on an exploratory basis.

To investigate these questions, we first collected a large sample of photos of faces that naturalistically varied in facial attractiveness and apparent intelligence. Separate samples of participants then either rated the male or female targets' attractiveness, intelligence, or humanness. We used Kteily et al.'s (2015) blatant dehumanization scale to evaluate humanness—a well-validated measure that robustly predicts important real-world social judgments between groups (e.g., Kteily et al., 2015) and individuals (e.g., Deska et al., 2018). The blatant dehumanization measure may have some advantages over other commonly-employed measures of dehumanization in this work because it does not conflate the particular dependent and independent variables. For instance, both the Mind Attribution Scale (e.g., measuring impressions of targets' ability to engage in higher-order thought; Kozak et al., 2006) and Haslam's "human uniqueness" traits (e.g., cognitive capacity; Haslam, 2006) measure higher-order mental capacities as signals of humanness, which would clearly confound treating variations in facial intelligence as an independent variable. With respect to the infrahumanization scale, its effects are "contingent on the existence of meaningful in-group/out-group distinctions," also rendering it unsuitable for investigating interpersonal dehumanization (Haslam & Loughnan, 2014). Importantly, we also followed earlier work by adjusting for targets' likeability in our analyses so as to distinguish the results from valence and related halo effects (Kteily et al., 2015; see also Bruneau et al., 2018).

To summarize, Study 1's design allowed for target-level tests of the key hypotheses that more attractive and more intelligent faces seem more human, and for exploring whether these hypothesized trait-linked associations with humanness manifest differently in male and female targets. We report all measures, manipulations, and exclusions in every study reported below. In additional analyses, we bootstrapped 1,000 resamples of all regression coefficients of interest to

obtain confidence intervals without assuming normality (e.g., Pardoe, 2000); results remained identical across the canonical and bootstrapped analyses.

Method

We recruited 412 undergraduate students participating for partial course credit to serve as targets (206 female, 206 male; M_{age} = 19.46 years, SD = 2.56; all White; 30 individuals did not report their age), thus achieving approximately 97% power in a multiple regression model with five predictors in a target-level analysis when assuming the average effect size in social and personality psychology (r = .21; Richard et al., 2003). After providing consent to use their images in subsequent research, we instructed the targets to pose with a neutral facial expression while photographing every participant against the same canvas paper background at a fixed distance and viewing angle from a tripod-mounted camera using a box lamp flash without any additional ambient illumination. We then cropped the photos to the edges of their head (including their hair), converted the images to grayscale, and resized them to a fixed height before presenting them to perceivers.

We next recruited 240 participants from Amazon's Mechanical Turk (MTurk) to rate the target faces. After completing informed consent, participants rated all 206 of the male or female targets on one of attractiveness (1 = *Not at all attractive*, 7 = *Extremely attractive*), perceived intelligence (1 = *Not at all intelligent*, 7 = *Extremely intelligent*), likeability (1 = *Not at all likeable*, 7 = *Extremely likeable*), or humanness (see below) in random order; approximately 30 perceivers made each judgment to ensure stable means (Hehman et al., 2018).

We used Kteily et al.'s (2015) Blatant Dehumanization scale to measure how human each target seemed. Participants were first instructed that "people can vary in how human-like they seem. Some people seem highly evolved whereas others seem no different from lower animals.

You will be asked to indicate how evolved you think each [pictured] individual is." We emphasized that there were no "right" or "wrong" answers and encouraged participants to rely on their "gut instinct" when responding. Participants responded to each face using a slider ranging from a graphic of *Dryopithcus* (a stooped, ancestral ape; i.e., 0% human) to an upright, walking *Homo sapiens* (i.e., 100% human; see Table 1 for descriptive statistics for the humanness ratings from Studies 1-4). Previous research established the convergent and divergent validity of this measure for indexing blatant dehumanization (Kteily et al., 2015). Finally, we thanked, debriefed, and compensated the participants.

Table 1

Descriptive Statistics for the Humanness Ratings Across Studies 1-4

Study	Targets	*N*	*M*	*SD*
1	All	412	80.48	5.38
	Female	206	83.22	4.23
	Male	206	77.75	5.01
2	All	412	84.37	3.36
	Female	206	85.94	3.12
	Male	206	82.81	2.84
3	All	412	80.94	4.94
	Female	206	84.31	3.17
	Male	206	77.58	4.02
4	All	88	76.25	3.95
	Female	44	77.52	3.90
	Male	44	74.97	3.60

Results

We first aggregated perceivers' judgments by calculating mean ratings of attractiveness, intelligence, likeability, and humanness for each face to enable a target-level analysis. Attractiveness, intelligence, likeability, gender, attractiveness × gender, and intelligence × gender served as predictors of the targets' humanness scores in a multiple regression. As expected, more attractive and intelligent-looking targets were rated as more evolved (Table 2). Attractiveness and perceived intelligence correlated for both male, $r(204) = .51$, $p < .001$, and female targets, $r(204) = .55$, $p < .001$, but did not demonstrate collinearity in any of the models (i.e., all VIFs < 7; Neter et al., 1990).

Table 2

Standardized Regression Coefficients and Test Statistics for the Model Predicting Targets' Judged Humanness From Their Attractiveness, Perceived Intelligence, Likeability, Gender, Attractiveness × Gender Interaction, and Perceived Intelligence × Gender Interaction

Predictor	β	*t*	*p*	$r_{\text{effect size}}$
Attractiveness	.41	9.31	< .001	.42
Perceived intelligence	.27	7.23	< .001	.34
Likeability	.15	2.62	.009	.13
Gender	-.15	-3.71	< .001	-.18
Attractiveness × Gender	-.16	-4.17	< .001	-.20
Perceived intelligence × Gender	.17	5.15	< .001	.25

Note. $N = 412$.

Gender coded -1 = female, 1 = male.

Notably, the importance of looking attractive and intelligent also differed depending on the target's gender. We therefore estimated multiple regression models to estimate the results

separately for the male and female targets by dummy coding the hypothesis-salient gender as 0 and the other gender as 1.

Recapitulating the interactions with gender above, facial attractiveness predicted women's humanness scores nearly two times better than men's, and facial intelligence predicted men's humanness scores nearly four times better than women's (Table 3). Moreover, attractiveness predicted women's humanness approximately five times better than perceived intelligence did, whereas perceived intelligence predicted men's humanness approximately twice as well as attractiveness did. Meta-analytic comparisons of the effect sizes showed that all of these associations significantly differed (Zs ≥ 3.41, ps $< .001$). All of the results above emerged independent of targets' likeability.

Table 3

Standardized Regression Coefficients and Test Statistics for Separate Multiple Regression Models Predicting Female and Male Targets' Judged Humanness From Their Attractiveness, Perceived Intelligence, and Likeability

Predictor	β	*t*	*p*	$r_{\text{effect size}}$
Female targets				
Attractiveness	.59	10.16	< .001	.45
Perceived intelligence	.11	2.06	.04	.10
Likeability	.10	1.44	.15	.07
Male targets				
Attractiveness	.22	3.24	.001	.16
Perceived intelligence	.42	8.30	< .001	.38
Likeability	.23	2.46	.01	.12

Note. $N = 412$.

Discussion

Overall, the data supported our hypotheses: Face-based trait impressions influenced humanness judgments such that less attractive and less intelligent-looking individuals seemed less human. Our exploratory analyses also revealed that target gender qualified the associations between perceptions of the traits and humanness independent of likeability: Looking attractive better facilitated ascriptions of humanness to women than to men, whereas looking intelligent better facilitated ascriptions of humanness to men than to women.

The fact that these trait-to-humanness links are gendered comports with gender stereotypes; society rewards attractive women and intelligent men (e.g., Smiler & Epstein, 2010). However, the exploratory nature of these tests necessitates replication with a new sample of perceivers. We did this in Study 2 while simultaneously extending the investigation to consider whether perceivers' endorsement of gender stereotypes might explain the gender-based difference in humanness ascriptions. Specifically, we expected that people who value attractiveness in women and intelligence in men would ascribe more humanness to targets who match those expectations compared to those who do not (i.e., unattractive women and unintelligent-looking men).

Study 2

In Study 1, we moved beyond work on group-based ascriptions of humanness to individual-based ascriptions. We found that within-group variability in targets' facial appearance promotes dehumanization such that less attractive and less intelligent-looking individuals were rated as less evolved. Moreover, attractiveness more strongly predicted women's humanness whereas perceived intelligence more strongly predicted men's humanness. These gender

differences accord with stereotypes that promote the importance of beauty in women and of agency-related traits (e.g., intelligence) in men (e.g., Smiler & Epstein, 2010).

Studies 2 and 3 were designed to replicate and extend these effects. In each of these studies, we again sought to directly replicate both that facial attractiveness and apparent intelligence are related to dehumanization. We also sought to replicate the gendered nature of this effect, expecting that a new sample of perceivers would demonstrate the same gender difference (see https://osf.io/x7e4z/?view_only=4abd538ad5c54849890a084de7087e33 for Study 2).[1] But in Studies 2 and 3 we also added measures to test whether the gender difference in humanness ascriptions may vary according to the participants' endorsement of gender stereotypes related to attractiveness and intelligence. Here in Study 2, we created a new measure of gender stereotype endorsement (described below), and in Study 3 we adapted a previous measure of value rankings for partner traits (see Buss, 1989), and conducted a within-paper meta-analysis of these findings across these studies (presented in the Study 3 Results). Of interest across Studies 2 and 3 was whether, across a meta-analysis of the data, we observed that people who endorsed valuing attractiveness in women and valuing intelligence in men would show stronger gendered effects. To measure the interaction between perceiver (endorsement of gender stereotypes) and target characteristics (attractiveness, perceived intelligence, and gender), we estimated the results using a cross-classified model (Westfall et al., 2014).

Method

[1] After receiving reviewer feedback, we deviated from our preregistered analysis plan by adding likeability ratings and using dummy coding to decompose the gender interactions. Moreover, models with full random slopes did not converge; we therefore estimated random-intercept models across Studies 2-5. Otherwise, all analyses, designs, and hypotheses in Studies 2-5 were preregistered unless stated in the main text.

We recruited a new sample of 256 participants from MTurk to achieve approximately 95% power in a cross-classified model with 412 targets and five predictors (Westfall et al., 2014). We randomly assigned the participants to rate the 206 male (N = 96; 69 female, 26 male, 1 "other;" M_{age} = 38.77 years, SD = 12.62; 77 White, 5 Black, 5 Hispanic, 5 "other," 3 East Asian, and 1 South Asian) or 206 female targets from Study 1 (N = 111; 69 female, 41 male, 1 "other;" M_{age} = 39.28 years, SD = 12.35; 88 White, 10 Hispanic, 6 Black, 3 East Asian, 3 South Asian, and 1 "other") in random order on the same dehumanization measure employed in Study 1.

Participants then completed 14 items assessing the extent to which they endorse gender-based stereotypes regarding attractiveness and intelligence respective to the gender of the targets that they judged (i.e., stereotypes of men for participants who rated men's faces, stereotypes of women for participants who rated women's faces). We constructed these items ourselves because we found no existing scales that precisely measure gendered beliefs about attractiveness and intelligence. Employing an agency-communion approach (Bakan, 1966), Items 1-6 assessed the importance of intelligence and attractiveness for predicting social success (a stereotypically feminine accomplishment), Items 7-12 assessed the importance of intelligence and attractiveness for predicting career success (a stereotypically masculine accomplishment; e.g., Smiler & Epstein, 2010), and the two final items measured the relative importance of each trait for members of the respective gender. Participants responded to all items using a 1 (*Strongly disagree*) to 7 (*Strongly agree*) scale (Appendix).

We planned to aggregate these items according to the results of a factor analysis and with consideration of our a priori expectations (i.e., that Items 1-3 and 4-6 would independently cluster to measure the importance of intelligence and attractiveness for communion, respectively;

that Items 7-9 and 10-12 would independently cluster to measure the importance of intelligence and attractiveness for agency, respectively; and that Items 13 and 14 would cluster to measure the general weighting of attractiveness and intelligence). We also predicted that participants' endorsement of these stereotypes would moderate how strongly they associate attractiveness and intelligence with humanness for men and women.

Results

We first excluded data from 49 participants who failed an attention check item ("In order for a man (woman) to be professional, I think it is important that (s)he is intelligent. Please do not respond to this question. In other words, skip it without selecting an answer;" Oppenheimer et al., 2009).

Preliminary analysis of gender-stereotyping items. Parallel analyses of Monte Carlo simulations (Horn, 1965) suggested retaining three factors for the items in both versions of the scale (i.e., male and female). Assuming oblique factors, we subsequently performed an exploratory factor analysis with Promax rotation specifying three factors to ascertain the item loadings. Using both approaches helped us to determine the appropriate factors and how many items loaded on each.

For both sets of items (i.e., with men and women as the named target), Items 1 and 2 (both measuring the importance of intelligence for communion), Items 7 and 8 (both measuring the importance of intelligence for agency), and Item 14 (prioritization of attractiveness over intelligence) loaded onto the same factor (all $\lambda s \geq .65$; all nonfactor $|\lambda s| \leq .38$). Additionally, Items 4 and 5 (both measuring the importance of attractiveness for communion), and Items 10 and 11 (both measuring the importance of attractiveness for agency) loaded onto another factor (all $\lambda s \geq .53$; all nonfactor $|\lambda s| \leq .37$). None of the reverse-scored items (3, 6, 9, and 12) loaded

onto these two factors (see Gehlbach & Brinkworth, 2011). Three of them formed a third factor for ratings of men (Items 3, 6, and 12, with $\lambda s \geq .45$; all nonfactor $|\lambda s| \leq .31$) and all four formed a third factor for ratings of women (all $\lambda s \geq .67$; all nonfactor $|\lambda s| \leq .17$). Finally, Item 13 (prioritization of intelligence over attractiveness) did not load onto any factors (all $|\lambda s| \leq .38$).

We therefore combined Items 1, 2, 7, and 8 into a Value of Intelligence in Men's (Women's) Lives subscale, and Items 4, 5, 10, and 11 into a Value of Attractiveness in Men's (Women's) Lives subscale, explaining 40% of the variance in the ratings for a male target and 43% of the variance in the ratings for a female target. Although we initially expected to add the reverse-scored items to their expected factors, they did not load as expected and therefore unacceptably reduced interitem reliabilities (including the reverse-scored items in the analyses nevertheless does not change the magnitude or significance of the reported results). Given that the two versions of Item 13 did not load robustly onto either factor and did not conform to our prior expectations, we also excluded it (and Item 14) from further analysis. When combined, the items for Value of Intelligence in Men's (Women's) Lives and Value of Attractiveness in Men's (Women's) Lives showed good interitem reliabilities (αs = .84, .80, .87, and .89, respectively) and only modestly correlated, $r(205) = .33$, $p < .001$.

Primary analyses. We first tested whether the results of Study 1 would replicate in this new sample. The targets' group-mean-centered attractiveness, perceived intelligence, and likeability (using the consensus scores from Study 1), gender, interaction between gender and attractiveness, and interaction between gender and perceived intelligence predicted the humanness ratings in a cross-classified model: Less attractive and less intelligent-looking individuals seemed less human, with attractiveness predicting women's humanness better than men's and perceived intelligence predicting men's humanness better than women's (Table 4).

We next used cross-classified models to test whether participants' gender stereotype endorsement moderated these associations, employing dummy coding to estimate effects within the hypothesis-salient gender.

Table 4

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Targets' Judged Humanness From Their Attractiveness, Perceived Intelligence, Likeability, Gender, Attractiveness × Gender Interaction, and Perceived Intelligence × Gender Interaction

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
All targets						
Attractiveness	2.36	0.17	399.80	13.79	< .001	.57
Perceived intelligence	2.89	0.29	396.50	10.14	< .001	.45
Likeability	1.24	0.31	390.60	3.96	< .001	.20
Gender	-0.94	1.13	206.20	-0.83	.41	-.06
Attractiveness × Gender	-0.52	0.15	403	-3.51	< .001	-.17
Perceived intelligence × Gender	0.88	0.26	397.90	3.41	< .001	.17
Female targets						
Attractiveness	3.05	0.22	372.50	13.89	< .001	.58
Perceived intelligence	2.30	0.41	372.50	5.63	< .001	.28
Likeability	0.64	0.38	372.50	1.68	.09	.09
Male targets						
Attractiveness	1.53	0.27	426.10	5.78	< .001	.27
Perceived intelligence	3.39	0.39	426.10	8.61	< .001	.38
Likeability	2.39	0.53	426.10	4.50	< .001	.21

Female targets. Attractiveness and perceived intelligence both predicted how human women seemed (independent of likeability), though attractiveness played a stronger role in judgments of women's humanness than men's (Tables 4-5). Several results indicated that attractiveness held more importance than perceived intelligence in predicting women's humanness because the participants believed that attractiveness matters more for women than intelligence does.

Table 5

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Female Targets' Judged Humanness from their Attractiveness, Perceived Intelligence, Likeability, Participants' Endorsement of Gender-Based Attractiveness and Intelligence Stereotypes, and Their Interactions

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Target predictors						
Attractiveness	3.10	0.22	374	14.09	< .001	.59
Perceived intelligence	2.37	0.41	374	5.78	< .001	.28
Likeability	0.64	0.38	373	1.68	.09	.09
Participant predictors						
Attractiveness stereotype endorsement	-3.74	1.05	201	-3.56	< .001	-.24
Intelligence stereotype endorsement	0.34	1.25	201	0.27	.79	.02
Target × participant predictors						
Attractiveness × Attractiveness stereotype endorsement	0.48	0.09	42010	5.21	< .001	.03
Attractiveness × Intelligence stereotype endorsement	-0.01	0.11	42010	-0.05	.96	.00

Perceived intelligence × Attractiveness stereotype endorsement	0.68	0.18	42010	3.86	< .001	.02
Perceived intelligence × Intelligence stereotype endorsement	-0.06	0.21	42010	-0.31	.76	.00

First, the more a participant valued attractiveness in women's lives, the more that person tended to rate women as generally less human. Second (and in contrast), the value participants placed on intelligence in women's lives did not relate to their humanness judgments. Thus, people held higher standards for women's humanness to the extent that they valued their attractiveness but not to the extent that they valued their intelligence. More simply: People who subscribed to the stereotype equating beauty and value for women dehumanized women more.

Third, simple slopes analyses (Aiken & West, 1991) showed that women's attractiveness more strongly influenced judgments of their humanness among participants who prioritized attractiveness in women, $b = 3.77$, $SE = 0.26$, $t(716.40) = 14.58$, $p < .001$, $r_{\text{effect size}} = .48$, compared to participants who did not, $b = 2.43$, $SE = 0.25$, $t(638.70) = 9.65$, $p < .001$, $r_{\text{effect size}} = .36$, even though attractiveness still predicted humanness judgments for the latter group.

Fourth, women's perceived intelligence influenced judgments of their humanness almost equally for participants who highly valued intelligence in women, $b = 2.28$, $SE = 0.49$, $t(780.50) = 4.64$, $p < .001$, $r_{\text{effect size}} = .16$, and for participants who did not highly value women's intelligence, $b = 2.45$, $SE = 0.48$, $t(732.50) = 5.05$, $p < .001$, $r_{\text{effect size}} = .18$. In other words, participants who reported highly valuing women's intelligence did not reflect this in their humanness judgments of women.

Male targets. Attractiveness and perceived intelligence both predicted how humanlike men seemed (independent of likeability). Perceived intelligence played a greater role for men

than it did for women (Tables 4 and 6). Several results indicated that participants especially focused on men's perceived intelligence when judging their humanness because they believed that intelligence matters more for men than attractiveness does.

Table 6

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Male Targets' Judged Humanness from their Attractiveness, Perceived Intelligence, Likeability, Participants' Endorsement of Gender-Based Attractiveness and Intelligence Stereotypes, and Their Interaction

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Target predictors						
Attractiveness	1.42	0.27	428	5.37	< .001	.25
Perceived intelligence	3.30	0.39	429	8.36	< .001	.37
Likeability	2.39	0.53	426	4.50	< .001	.21
Participant predictors						
Attractiveness stereotype endorsement	-1.28	1.25	201	-1.02	.31	-.07
Intelligence stereotype endorsement	-3.91	1.27	201	-3.08	.002	-.21
Target × participant predictors						
Attractiveness × Attractiveness stereotype endorsement	0.45	0.12	42,010	3.65	< .001	.02
Attractiveness × Intelligence stereotype endorsement	0.80	0.13	42,010	6.42	< .001	.03
Perceived intelligence ×Attractiveness stereotype endorsement	0.20	0.20	42,010	1.04	.30	.00
Perceived intelligence × Intelligence	1.00	0.20	42,010	5.05	< .001	.02

stereotype endorsement

First, the more a participant valued intelligence in men's lives, the more that person tended to rate men as generally less human. Second (and in contrast), participants who valued attractiveness in men's lives did not judge their humanness more harshly. Thus, people held higher standards for men's humanness according to how much they valued their intelligence but not their attractiveness. More simply: People who subscribed to the stereotype equating intelligence and value for men dehumanized them more.

Third, targets' intelligence more strongly influenced judgments of men's humanness among participants who prioritized intelligence in men, $b = 4.62$, $SE = 0.51$, $t(371) = 9.01$, $p < .001$, $r_{\text{effect size}} = .42$, compared to participants who did not, $b = 1.98$, $SE = 0.52$, $t(388) = 3.81$, $p < .001$, $r_{\text{effect size}} = .19$, even though perceived intelligence still predicted humanness judgments for the latter group.

Exploratory comparisons of gender-stereotype endorsement. Thus far, the results suggest that gender stereotype endorsement can help to explain why attractiveness and perceived intelligence differently influence judgments of men's and women's humanness. Yet, we also found that attractiveness influenced humanness ratings among participants who reported highly valuing men's attractiveness, $b = 3.57$, $SE = 0.51$, $t(353) = 7.05$, $p < .001$, $r_{\text{effect size}} = .35$, compared to participants who did not, $b = 0.82$, $SE = 0.35$, $t(380) = 2.37$, $p = .02$, $r_{\text{effect size}} = .12$. This may indicate that beliefs about the importance of attractiveness similarly influence humanness judgments for both men and women. In other words, gender stereotypes may not explain why attractiveness affects judgments of women's humanness more than men's.

Two exploratory comparisons helped us to probe this question and further substantiate our conclusion regarding these gender differences. First, we created a model with targets'

gender; group-mean-centered attractiveness, perceived intelligence, and likeability (using the consensus scores from Study 1); and interactions between target gender, target attractiveness/perceived intelligence, and participants' stereotype endorsement (i.e., target gender × target attractiveness × attractiveness stereotype endorsement, target gender × target attractiveness × intelligence stereotype endorsement, target gender × target perceived intelligence × attractiveness stereotype endorsement, target gender × target perceived intelligence × intelligence stereotype endorsement) to predict the humanness ratings in a cross-classified model. The three-way interaction terms (e.g., target gender × target perceived intelligence × intelligence stereotype endorsement) test whether the difference in influence between those who value a trait highly and those who do not value a trait is greater for ratings of one target gender over the other.

To illustrate, imagine that (when rating female targets' humanness) people who highly value women's intelligence are influenced by the female targets' apparent intelligence twice as much as people who do not value women's intelligence. Further imagine that (when rating male targets' humanness) people who highly value men's intelligence are influenced by the male targets' apparent intelligence four times as much as people who do not value men's intelligence. In this hypothetical case, targets' perceived intelligence interacts with participants' endorsement of the intelligence stereotype for ratings of both target genders, indicating that the degree to which a participant values others' intelligence, whether male or female, influences how much they regard (un)intelligent-looking looking men and women as human. This may seem to indicate that the endorsement of the intelligence stereotype does not account for why male targets' humanness depends more on their apparent intelligence compared to female targets. Yet, it does actually account for why men's humanness depends more on their perceived intelligence

than women's humanness does because the difference between people who highly value (men's/women's) intelligence and people who do not highly value intelligence is twice as large for the male versus female targets (i.e., four times as much for male targets versus two times as much for female targets). Thus, the *relative* magnitude of the importance of the trait differs between judgments of men and women.

In assessing relative magnitude, the three-way interaction for attractiveness was not significant (i.e., target gender × target attractiveness × attractiveness stereotype endorsement), $b = -0.03$, $SE = 0.15$, $t(42010) = -0.19$, $p = .85$, $r_{\text{effect size}} = .00$, whereas the three-way interaction for intelligence was significant (i.e., target gender × target perceived intelligence × intelligence stereotype endorsement), $b = 1.06$, $SE = 0.29$, $t(42010) = 3.69$, $p < .001$, $r_{\text{effect size}} = .02$. Thus, the relative difference in humanness ratings between participants who highly value attractiveness and participants who do not highly value attractiveness was similar (i.e., not significant) for their ratings of male and female targets. Conversely, the relative difference in humanness ratings between participants who highly value intelligence and participants who do not highly value intelligence was greater for male targets than for female targets.

These relative magnitude analyses help to further substantiate that intelligence stereotype endorsement accounts for why perceived intelligence influences ratings of male targets' humanness more than female targets' humanness. However, they do not explain why attractiveness influences ratings of female targets' humanness more than male targets' humanness. Because the analyses only compare relative differences between genders among participants who highly value a trait versus participants who do not highly value that trait, the possibility that participants who highly value a trait are more influenced by it in *absolute* terms for one gender versus another gender remains open. In other words, participants who highly

value attractiveness likely value it in both men and women, but also likely more so in women. Conversely, participants who do not highly value attractiveness likely do not particularly value it in either men or women, but still more so in women. Analogous differences should apply to intelligence.

To further illustrate: imagine that the standardized regression coefficient for how much target attractiveness predicts ratings of male targets' humanness among people who highly value men's attractiveness is $\beta = .25$, whereas it is $\beta = .10$ for people who do not value men's attractiveness. Further imagine that the standardized regression coefficient for how much target attractiveness predicts ratings of female targets' humanness among people who highly value women's attractiveness is $\beta = .50$, whereas it is $\beta = .20$ among people who do not value women's attractiveness. In each case, endorsement of the attractiveness stereotype influences the degree to which men's and women's attractiveness affects ratings of their humanness (because the standardized regression coefficients are larger in magnitude among people who highly value attractiveness vs. people who do not for both genders; i.e., .25 vs. .10 for male targets, and .50 vs. .20 for female targets). Indeed, the relative magnitude is multiplicatively equal (.25 is 2.5 times larger than .10, and .50 is 2.5 times larger than .20). But the absolute magnitude for the female targets is double that for the male targets (i.e., .50 vs. .25, and .20 vs. .10). Thus, absolute differences in effect sizes can also signal differences in the impact of stereotype endorsement across genders. To better capture these absolute differences, we conducted a second series of exploratory analyses comparing the effect sizes that measure how strongly each trait predicts target gender differences among participants who highly value the trait and among participants who do not highly value the trait (i.e., the effect sizes ascertained from the simple slopes analyses of the target × participant interaction terms from the models outlined in Tables 5 and 6).

Among participants who highly value attractiveness, women's attractiveness ($r_{\text{effect size}}$ = .60) influenced their humanness judgments significantly more than men's attractiveness did ($r_{\text{effect size}}$ = .35), meta-analytic $Z = 4.62$, $p < .001$. Similarly, among participants who do not highly value attractiveness, women's attractiveness ($r_{\text{effect size}}$ = .46) still influenced their humanness judgments more than men's attractiveness did ($r_{\text{effect size}}$ = .12), meta-analytic $Z = 5.25$, $p < .001$.

Conversely, among participants who highly value intelligence, men's perceived intelligence ($r_{\text{effect size}}$ = .42) influenced their humanness judgments significantly more than women's perceived intelligence did ($r_{\text{effect size}}$ = .23), meta-analytic $Z = 3.06$, $p = .002$. Finally, among participants who do not highly value intelligence, men's perceived intelligence ($r_{\text{effect size}}$ = .19) did not influence humanness judgments more than women's perceived intelligence did ($r_{\text{effect size}}$ = .26), meta-analytic $Z = 1.05$, $p = .29$. Triangulating across the preregistered and exploratory analyses, these results indicate that gender stereotypes help to account for why attractiveness better predicted judgments of women's humanness and why perceived intelligence better predicted judgments of men's humanness.

Exploratory analyses of participant gender effects. Finally, we examined whether participant gender affects the results of the main analyses described above. To do so, we added participant gender as a predictor interacting with target gender and target trait (intelligence, attractiveness) to predict targets' humanness in the cross-classified models. We then used dummy coding to further explore significant interactions involving participant gender (0 = hypothesis-salient gender, 1 = other gender). Indeed, we observed three-way interactions between participant gender, target gender, and trait for both attractiveness, $b = 1.28$, $SE = 0.44$, $t(41602) = 2.92$, $p = .003$, $r_{\text{effect size}}$ = .01, and intelligence: $b = 2.25$, $SE = 0.75$, $t(41602) = 3.01$, p

= .003, $r_{\text{effect size}}$ = .01. We decompose these interactions by target gender using dummy coding below. We did not observe main effects for participant gender and thus do not discuss them further.

Female targets. Within female targets, we observed an interaction between participant gender and target attractiveness, b = -1.21, SE = 0.27, t(41602) = -4.56, p < .001, $r_{\text{effect size}}$ = .02. Specifically, female raters, b = 3.36, SE = 0.24, t(555) = 14.22, p < .001, $r_{\text{effect size}}$ = .52, used female targets' attractiveness more than male raters did, b = 2.15, SE = 0.27, t(970) = 7.90, p < .001, $r_{\text{effect size}}$ = .25, when rating female targets' humanness. Conversely, female raters, b = 1.94, SE = 0.44, t(562) = 4.36, p < .001, $r_{\text{effect size}}$ = .18, used perceived intelligence about as much as male raters did, b = 2.20, SE = 0.51, t(999) = 4.28, p < .001, $r_{\text{effect size}}$ = .13, when rating female targets' humanness.

Male targets. Within male targets, we observed an interaction between participant gender and targets' perceived intelligence, b = 2.51, SE = 0.55, t(41604) = 4.59, p < .001, $r_{\text{effect size}}$ = .02. Specifically, male raters, b = 5.61, SE = 0.54, t(1845) = 10.35, p < .001, $r_{\text{effect size}}$ = .23, used male targets' perceived intelligence more than female raters did, b = 3.10, SE = 0.41, t(583) = 7.65, p < .001, $r_{\text{effect size}}$ = .30, when rating male targets' humanness. Conversely, both male, b = 1.91, SE = 0.35, t(1745) = 5.50, p < .001, $r_{\text{effect size}}$ = .13, and female raters, b = 1.84, SE = 0.26, t(572) = 7.05, p < .001, $r_{\text{effect size}}$ = .28, used attractiveness similarly when rating male targets' humanness.

Discussion

In a preregistered replication of Study 1, we again observed that unattractive and unintelligent-looking individuals seemed less human than their attractive and intelligent-looking counterparts. We also again observed that attractiveness better predicted judgments of women's humanness than men's, whereas perceived intelligence better predicted judgments of men's

humanness than women's (independent of likeability). Furthermore, participants' endorsement of gender stereotypes regarding male intelligence and female beauty predicted this emphasis on women's attractiveness and men's perceived intelligence when evaluating their humanness.

Including participant gender in the analyses largely demonstrated the same pattern of results as before, but also added interesting nuance. Specifically, women considered women's attractiveness more than men did when rating women's humanness, whereas men considered men's apparent intelligence more than women did when rating men's humanness. These findings recall notions of intragender competition: Women more harshly derogate other women based on their attractiveness to gain advantages, whereas men more harshly derogate other men based on their competence to gain advantages (e.g., Buss, 1988, 1994; Buss & Dedden, 1990). Future research should therefore explore intragender competition as an arena in which dehumanization may be rife.

Consistent with predictions, women's attractiveness especially influenced judgments of their humanness among participants who prioritized attractiveness in women, and men's apparent intelligence especially influenced judgments of their humanness among participants who prioritized intelligence in men. Notwithstanding these and other theoretically consistent results (see above), two of the interaction terms in the target gender models (Tables 5-6) challenged the precision of our stereotype endorsement scale, thus demonstrating some inconsistency with the notion that the valuing of women's beauty and men's intelligence leads to their differential dehumanization. Specifically, participants who valued women's attractiveness were more influenced by women's apparent intelligence than were participants who do not value women's attractiveness, and participants who valued men's intelligence were more influenced by men's attractiveness than were participants who do not value men's intelligence. We suspect that

such inconsistent results may have emerged because we needed to create a measure of gender-based valuing of intelligence and beauty rather than relying on a well-validated measure; some noise is to be expected. Thus, we employed another measure of valuing beauty and intelligence in Study 3 and then meta-analyzed the results across these two studies.

Study 3

In Study 2, we found some evidence that women and men are differently dehumanized due to their apparent intelligence and attractiveness because of gender stereotypes, but also observed some surprising results. Consistent with our logic above, we conducted a preregistered conceptual replication of Study 2 using a different measure of valuing intelligence and attractiveness and meta-analyzed the results across Studies 2 and 3 (see https://osf.io/p3r8n/?view_only=7a03a6a3a9a5484eb1efa738b3041fca).

In Study 3, we adapted a ranking measure used in previous research examining the importance of various traits in the desirability of men and women as romantic partners (we removed the reference to desirability as a partner; e.g., Buss, 1989). We expected that participants who ranked men's intelligence (women's beauty) as higher in importance for men (women) to possess would be especially influenced by the male targets' apparent intelligence (female targets' attractiveness) when rating their humanness.

Method

We recruited 305 participants from MTurk to achieve approximately 95% power in a cross-classified model with 412 targets (Westfall et al., 2014) and to account for expected exclusions due to participant or survey-loading errors. We randomly assigned the participants to rate either the 206 male (N = 95; 53 female, 42 male; M_{age} = 40.96 years, SD = 13.23; 71 White, 8 East Asian, 7 Black, 5 "other," 2 Hispanic, and 2 South Asian) or 206 female targets from

Study 1 (N = 107; 54 female, 52 male, 1 "other;" M_{age} = 43.68 years, SD = 14.17; 78 White, 15 Black, 5 Hispanic, 4 "other," 3 East Asian, and 2 South Asian) in random order on the same dehumanization measure employed in Study 1. We excluded data from 80 participants for failing the attention check (same as in Study 2), 18 participants for reporting that the face photos did not load, 3 participants for reporting that they randomly ranked the traits, and 2 participants for reporting that they randomly rated the faces.

Participants then ranked 12 traits on their importance for people of the same gender as the targets they judged on humanness:

> Below is a list of 12 traits that could describe an individual. We want you to think about how important each trait is for men [women] to possess. For example, is it important for men [women] to be assertive? Enthusiastic? Etc. Do not think about a specific man [woman]. Instead, think about how important each trait is for men [women] in general. Please rank the traits on how important they are for men [women] to possess from 1 (most important) to 12 (least important).

To broadly encapsulate the various traits that could describe an individual, we included two items for each of the Big Five traits (*assertive*, *enthusiastic*, *hard-working*, *organized*, *open to new experiences*, *creative*, *kind*, *polite*, *calm*, and *positive*), and added *intelligent* and *physically attractive* as our focal items. For consistency with Study 2, we describe participants' rankings of the importance of attractiveness and intelligence for men and women as their gender stereotype endorsement. Importantly, Study 3 differs from Study 2 in that higher scores on the gender stereotype endorsement measure indicate less endorsement (because a larger number indicates a lower rank; e.g., ranking intelligence as 12th in importance is a larger number, but

lower in rank, than ranking it as 1st in importance), whereas in Study 2 it indicated more endorsement. We maintained this direction when analyzing the Study 3 data but reversed the sign of effect sizes from Study 3 for the meta-analysis to enable aggregation with the effect sizes from Study 2.

Results

We first tested whether the results of Studies 1 and 2 replicated in this new sample. The targets' gender; group-mean-centered attractiveness, perceived intelligence, and likeability (using the consensus scores from Study 1); interaction between gender and attractiveness; and interaction between gender and perceived intelligence predicted the humanness ratings in a cross-classified model. As before, less attractive and less intelligent-looking individuals seemed less human, and perceived intelligence predicted men's humanness better than women's humanness (independent of likeability; Table 7). We did not find that attractiveness predicted women's humanness significantly better than men's, however (i.e., the attractiveness × gender interaction term was not significant). We next used cross-classified models to test whether participants' gender stereotype endorsement moderated these associations, employing dummy coding to estimate effects within the hypothesis-salient gender.

Table 7

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Targets' Judged Humanness From Their Attractiveness, Perceived Intelligence, Likeability, Gender, Attractiveness × Gender Interaction, and Perceived Intelligence × Gender Interaction

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$

All targets						
Attractiveness	2.33	.25	382.88	9.23	< .001	.43
Perceived intelligence	2.01	.47	382.46	4.25	< .001	.21
Likeability	1.80	.37	392.85	4.86	< .001	.24
Gender	-5.47	2.51	205.88	-2.18	.03	-.15
Attractiveness × Gender	.22	.35	401.04	0.63	.53	.04
Perceived intelligence × Gender	2.09	.61	397.52	3.44	< .001	.17
Female targets						
Attractiveness	2.59	.26	379.26	9.98	< .001	.46
Perceived intelligence	2.45	.48	379.26	5.07	< .001	.25
Likeability	0.87	.45	379.26	1.92	.06	.10
Male targets						
Attractiveness	2.07	.31	416.61	6.68	< .001	.31
Perceived intelligence	3.52	.46	416.61	7.63	< .001	.35
Likeability	3.57	.62	416.61	5.74	< .001	.27

Target gender: Male = 1, Female = -1.

Female targets. Attractiveness and perceived intelligence both predicted how human women seemed (independent of likeability; Table 8). Replicating Study 2, the more a participant valued attractiveness in women, the more that person tended to rate women as generally less human, whereas the same pattern did not emerge for intelligence stereotype endorsement.

Table 8

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Female Targets' Judged Humanness From Their Attractiveness, Perceived

Intelligence, Likeability, Participants' Endorsement of Gender-Based Attractiveness and Intelligence Stereotypes, and Their Interactions

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Target predictors						
Attractiveness	5.19	.45	3293	11.53	< .001	.20
Perceived intelligence	2.36	.85	3546	2.77	.01	.05
Likeability	0.87	.45	379.40	1.92	.06	.10
Participant predictors						
Attractiveness stereotype endorsement	.95	.45	196	2.09	.04	.15
Intelligence stereotype endorsement	-.98	.63	196	-1.57	.12	-.11
Target × participant predictors						
Attractiveness × Attractiveness stereotype endorsement	-.35	.04	40990	-9.04	< .001	-.04
Attractiveness × Intelligence stereotype endorsement	.13	.05	40990	2.49	.01	.01
Perceived intelligence × Attractiveness stereotype endorsement	.10	.07	40990	1.43	.15	.01
Perceived intelligence × Intelligence stereotype endorsement	-.27	.10	40990	-2.67	.008	-.01

Note. For stereotype endorsement, a higher magnitude indicates less endorsement (i.e., lower ranking of the trait).

Moreover, the interactions between female targets' apparent traits and participants' stereotype endorsement indicated a valuing of women's attractiveness to the exclusion of their intelligence (and vice versa). Namely, the more a participant valued women's attractiveness, the more strongly their ratings of the female targets' humanness were influenced by the female

targets' attractiveness. Relatedly, the more a participant valued women's intelligence, the more strongly the female targets' apparent intelligence influenced that person's ratings of the female targets' humanness. Conversely, the more a participant valued women's intelligence, the less that the female targets' attractiveness influenced their ratings of the female targets' humanness.

Male targets. Attractiveness and perceived intelligence both predicted how human men seemed (independent of likeability), though perceived intelligence predicted men's dehumanization more than women's (Tables 7 and 9). Replicating Study 2, the more a participant valued men's intelligence, the more that person's ratings of men's humanness were influenced by the male targets' apparent intelligence. Similarly, the more a participant valued men's attractiveness, the more that person's ratings of men's humanness were influenced by the male targets' attractiveness. Yet, the other two interaction terms (attractiveness × intelligence stereotype endorsement and perceived intelligence × attractiveness stereotype endorsement) were also significant, indicating that the more that intelligence (attractiveness) was valued in men, the more a participant's ratings of male targets' humanness was influenced by their attractiveness (apparent intelligence). Finally, greater endorsement of the attractiveness stereotype predicted rating male targets as generally less human. Thus, we again observed overlap between the importance of intelligence and attractiveness in influencing ratings of men's humanness.

Table 9

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Male Targets' Judged Humanness From Their Attractiveness, Perceived Intelligence, Likeability, Participants' Endorsement of Gender-Based Attractiveness and Intelligence Stereotypes, and Their Interaction

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Target predictors						
Attractiveness	5.96	.55	3914	10.84	< .001	.17
Perceived intelligence	5.86	.85	4590	6.87	< .001	.10
Likeability	3.57	.62	416.50	5.74	< .001	.27
Participant predictors						
Attractiveness stereotype endorsement	1.89	.52	196	3.68	< .001	.25
Intelligence stereotype endorsement	-.29	.64	196	-0.46	.65	-.03
Target × participant predictors						
Attractiveness × Attractiveness stereotype endorsement	-.37	.05	40,990	-7.53	< .001	-.04
Attractiveness × Intelligence stereotype endorsement	-.20	.06	40,990	-3.33	< .001	-.02
Perceived intelligence × Attractiveness Stereotype endorsement	-.18	.08	40,990	-2.29	.02	-.01
Perceived intelligence × Intelligence stereotype endorsement	-.21	.10	40,990	-2.16	.03	-.01

Note. For stereotype endorsement, a higher magnitude indicates less endorsement (i.e., lower ranking of the trait).

***Meta-Analysis*.** Given that neither of the gender stereotype endorsement measures we used have been validated for assessing stereotypes regarding the importance of men and women's intelligence and beauty (indeed, no such measure exists) and the inconsistencies in the results across Studies 2 and 3, we conducted a random effects meta-analysis to ascertain which effects remained robust upon combination of the two sets of results. For this purpose, we reversed the sign of the effect sizes from Study 3 so that they matched Study 2 (in Study 2, a higher

magnitude on the stereotype endorsement measure indicates greater endorsement whereas a higher magnitude indicates lower endorsement in Study 3 because it means a lower rank). We conducted separate random effects meta-analyses for the female and male targets.

For female targets, only two effects were significant in the meta-analysis (Table 10). First, the more that participants valued attractiveness in women, the more they tended to rate female targets as generally less human. Second, the more that participants valued women's attractiveness, the more that female targets' attractiveness predicted their rating female targets as more human. Notably, neither female targets' apparent intelligence nor the valuing of women's intelligence influenced ratings of the female targets' humanness in this meta-analysis.

For male targets, the more that participants valued men's attractiveness, the more that male targets' attractiveness predicted their rating the male targets as more human (Table 11). Similarly, the more that participants valued men's intelligence, the more that male targets' apparent intelligence predicted their rating the male targets as more human. Furthermore, the more that participants valued men's intelligence, the more that male targets' attractiveness predicted their rating the male targets as more human.

Finally, given that the interaction between target gender and attractiveness did not reach significance in Study 3, we meta-analyzed the target gender and trait interaction effects across the studies using adult targets (i.e., Studies 1, 2, 3, and 5). When aggregated, the attractiveness × gender interaction was significant, indicating that attractiveness better predicted judgments of women's humanness than men's, $Z = -3.26$, $p = .001$, $\bar{r} = -.18$. The perceived intelligence × gender interaction was also significant, indicating that perceived intelligence better predicted judgments of men's humanness than women's, $Z = 6.07$, $p < .001$, $\bar{r} = .18$.

Table 10

Random Effects Meta-Analysis of the Stereotype Endorsement Effects for Female Targets Across Studies 2 and 3

	Study 2			Study 3			Meta-Analysis		
Predictors	*df*	*p*	$r_{\text{effect size}}$	*df*	*p*	$r_{\text{effect size}}$	*Z*	*p*	Mean *r*
Participant predictors									
Attractiveness stereotype endorsement	201	< .001	-.24	196	.04	-.15	-3.50	< .001	-.20
Intelligence stereotype endorsement	201	.79	.02	196	.12	.11	1.16	.25	.06
Target × participant predictors									
Attractiveness × Attractiveness stereotype endorsement	42,010	< .001	.03	40,990	< .001	.04	7.02	< .001	.04
Attractiveness × Intelligence stereotype endorsement	42,010	.96	.00	40,990	.01	-.01	-1.00	.32	-.01
Perceived intelligence × Attractiveness stereotype endorsement	42,010	< .001	.02	40,990	.15	-.01	0.34	.73	.01
Perceived intelligence × Intelligence stereotype endorsement	42,010	.76	.00	40,990	.008	.01	1.00	.32	.01

Note. We reversed the sign of the effect sizes from Study 3 to match Study 2's effect size direction.

Table 11
Random Effects Meta-Analysis of the Stereotype Endorsement Effects for Male Targets Across Studies 2 and 3

	Study 2			Study 3			Meta-Analysis		
	df	*p*	$r_{\text{effect size}}$	*df*	*p*	$r_{\text{effect size}}$	*Z*	*p*	Mean *r*
Participant predictors									
Attractiveness stereotype endorsement	201	.31	-.07	196	< .001	-.25	-1.82	.07	-.16
Intelligence stereotype endorsement	201	.002	-.21	196	.65	.03	-0.79	.43	-.09
Target × participant predictors									
Attractiveness × Attractiveness stereotype endorsement	42,010	< .001	.02	40,990	< .001	.04	3.11	.002	.03
Attractiveness × Intelligence stereotype endorsement	42,010	< .001	.03	40,990	< .001	.02	5.03	< .001	.03
Perceived intelligence × Attractiveness stereotype endorsement	42,010	.30	.00	40,990	.02	.01	1.00	.32	.01
Perceived intelligence × Intelligence stereotype endorsement	42,010	< .001	.02	40,990	.03	.01	5.02	< .001	.03

Note. We reversed the sign of the effect sizes from Study 3 to match Study 2's effect size direction.

Discussion

Aiming to clarify why people dehumanize men and women differently depending on their apparent intelligence and attractiveness, we conducted a preregistered replication of Study 2 using a new measure of gender stereotype endorsement. We then meta-analytically combined the results with those from Study 2 to identify the consistent effects.

The new data showed that gender influenced participants' ratings of the female targets' humanness in a hydraulic fashion: When rating the female targets' humanness, participants either valued attractiveness exclusively, or valued the female targets' apparent intelligence to the exclusion of their attractiveness. Ratings of male targets were only somewhat consistent with our theory, however.

Apparent intelligence influenced humanness ratings for male targets more than for female targets. Although the interaction between target gender and attractiveness did not reach significance here, the effect was in the same direction as in Studies 1, 2, and 5 (where it was significant). Moreover, the more that participants valued men's intelligence, the more that male targets' apparent intelligence influenced ratings of their humanness. Curiously, the other target × participant predictor interactions indicated that male targets' attractiveness also influenced ratings of their humanness.

To clarify the results across Studies 2 and 3, we conducted a random effects meta-analysis to combine the two studies' results. First, our overall effects replicate clearly. Both attractiveness and perceived intelligence matter for inferences about others' humanness, but in gendered ways. Women are more strongly affected by facial attractiveness, men by apparent intelligence. Second, the effects of gendered stereotype endorsement help us understand these effects. The aggregate results for ratings of female targets' humanness clearly match predictions:

Ratings of female targets' humanness depended only on how attractive they were and how much participants valued attractiveness in women. The pattern for male targets was more nuanced. As predicted, ratings of male targets' humanness depended on how intelligent they looked and how much participants valued intelligence in men. A similar effect occurred for attractiveness stereotypes; people who valued men's attractiveness used variations in men's attractiveness more in their humanness ratings. Surprisingly though, the meta-analysis also demonstrated that participants who valued men's intelligence also dehumanized less attractive male targets.

Why might those who endorse valuing men's intelligence also rely on men's attractiveness in dehumanizing judgments? Although speculative, it is possible that this reflects how gender stereotypes influence how people differently judge intelligence and attractiveness from facial structural cues in men's and women's faces. Women's facial attractiveness strongly correlates with their degree of facial femininity (O'Toole et al., 1998; Rhodes et al., 2000), whereas many of the features associated with competence are stereotypically masculine and thus unattractive in a female face (Praino et al., 2014). Consistent with this, competence is regarded as unattractive for women (Horner, 1972), with women who display competence being evaluated as antagonistic, cold, and unlikeable, thereby eliciting negative affect in perceivers (Butler & Geis, 1990; Hagan & Kahn, 1975; Porter & Geis, 1981). Reciprocally, attractive women are commonly perceived as incompetent (Kalof, 1999). These patterns also parallel ambivalent sexism, whereby women exhibiting traditional femininity are positively evaluated, whereas non-traditional women, such as those exhibiting agency and competence, meet hostility. Indeed, a number of theorists have argued that representations of female competence conflict with representations of traditional femininity (e.g., Diekman, 2007; Fiske et al., 2002; Heilman,

2001), an effect that appears to occur for both stereotype representations and facial representations.

Conversely, the cues signaling attractiveness and competence in male faces show some overlap, such that some of the facial features associated with a competent appearance (i.e., masculine facial features) can enhance men's attractiveness (e.g., DeBruine et al., 2006). Correspondingly, whereas a women's attractiveness harms perceptions of her performance in tasks requiring competence, men's attractiveness either benefits perceptions of his performance or remains orthogonal (Heilman & Saruwatari, 1979; Heilman & Stopeck, 1985). Thus, ratings of men's humanness might have related more liberally to attractiveness and intelligence because these traits overlap for men (and clash for women).

Beyond these nuances, evidence across Studies 2 and 3 suggest that gender stereotypes regarding the importance of women's beauty and men's intelligence play a robust role in perceptions of their humanness. We return to this issue when considering future directions in the General Discussion.

Considering that the stereotypes governing such perceptions of men and women can originate early in life (potentially scaffolding the ways that boys and girls think of themselves as they grow into adulthood; e.g., Signorielli, 1990), we wondered whether adults' perceptions of children's attractiveness and intelligence might influence how human they seem. Previous research shows that adults favor pretty girls and smart boys, and that adults' perceptions of children can meaningfully influence their life outcomes (e.g., Jacobs, 1991; Rosenthal & Jacobson, 1968; Signorielli, 1990). We thus conducted Study 4 to investigate whether the association between humanness and individuals' perceived traits may become conditioned early in life such that attractive girls and intelligent boys garner higher ascriptions of humanness.

Study 4

The results of Studies 1-3 indicate that men's and women's perceived humanness depends on how intelligent and attractive they look. Specifically, men's perceived humanness especially relies on how intelligent they appear, and women's perceived humanness especially relies on their physical attractiveness. Moreover, perceivers' endorsement of gender stereotypes helps to explain these associations. Considering the pervasive and early nature by which gender stereotypes manifest, we therefore wondered whether these trait-contingent ascriptions of humanity extend to perceptions of children. If these biases do emerge when judging children, they could affect the traits that they regard as important, and to which they aspire as they mature.

Although past research has shown that people see children as less mentally sophisticated than adults (Gray et al., 2007), comparatively few studies have focused on the factors that might influence the dehumanization of children (cf. Goff et al., 2014). Here, we investigated whether the tendency to dehumanize unattractive women and unintelligent-looking men would extend to girls and boys. In doing so, we sought to illuminate the factors that may influence ascriptions of children's humanness and, most germane, to examine whether the disparate associations between perceived intelligence, attractiveness, and humanness affect judgments of targets early in their lives.

We thus replicated the original design of Study 1 using White children's neutral faces, preregistering both the hypothesis that we would observe the same results as in Studies 1-3, but also the concern that people might not comply with reporting (or considering) children's humanness (see https://osf.io/ftnvh/?view_only=f612a74dc9bc44ba9b72889b7940a5d2). We also preregistered our plan to conduct a participant-level analysis (in contrast to the target-level analyses above) so that we could achieve sufficient statistical power by recruiting enough

participants to compensate for the limited number of child targets available in existing face databases.

Method

We exhausted all available research databases of White children's neutral expression faces, resulting in facial photos of 44 boys and 44 girls (Egger et al., 2011; Langner et al., 2010; LoBue & Thrasher, 2015). Targets' ages ranged from 3 to 12 years (M_{age} = 6.69 years, *SD* = 2.49). We cropped the photos to the edges of their heads (including their hair), converted them to grayscale, and resized them to a fixed height before presenting them to 60 MTurk participants randomly assigned to rate all 88 children's attractiveness (1 = *Not at all attractive*, 7 = *Extremely attractive*) or perceived intelligence (1 = *Not at all intelligent*, 7 = *Extremely intelligent*).

We then recruited 501 new MTurk participants to rate the children's humanness. We randomly assigned the participants to rate all of the boys (*N* = 249; 146 female, 103 male; M_{age} = 39.66 years, *SD* = 13.16; 185 White, 17 Black, 16 East Asian, 15 Hispanic, 9 "other," and 7 South Asian) or girls (*N* = 252; 139 female, 113 male; M_{age} = 38.82 years, *SD* = 12.37; 194 White, 19 Black, 14 East Asian, 12 Hispanic, 7 "other," and 6 South Asian) in random order using the blatant dehumanization measure from above. Although we had preregistered a participant-level analysis, we later learned that this analysis has some key limitations in this context. We therefore only used target-level and cross-classified models to analyze our data (though using the original participant-level analysis returns identical results).

Results

In the overall model, attractiveness predicted how human children seemed, $\beta = .76$, $t(82) = 7.94$, $p < .001$, whereas perceived intelligence did not, $\beta = .10$, $t(82) = 0.71$, $p = .48$. The cross-classified model showed the same pattern: Attractiveness significantly predicted humanness, *b* =

5.33, $SE = 0.67$, $t(81.88) = 7.93$, $p < .001$, $r_{\text{effect size}} = .66$, whereas perceived intelligence did not, $b = 0.68$, $SE = 0.98$, $t(81.88) = 0.70$, $p = .49$, $r_{\text{effect size}} = .08$.

The interaction between targets' attractiveness and gender (dummy-coded: 0 = hypothesis-salient gender, 1 = other gender) did not approach significance in the target-level analysis, $\beta = -.08$, $t(82) = -0.43$, $p = .67$, suggesting that attractiveness similarly predicted humanness for both boys, $\beta = .68$, $t(82) = 4.33$, $p < .001$, and girls, $\beta = .76$, $t(82) = 7.94$, $p < .001$. The interaction between targets' perceived intelligence and gender also did not reach significance, $\beta = .19$, $t(82) = 1.09$, $p = .28$. Similarly, the cross-classified model did not yield significant gender-trait interaction effects for either attractiveness, $b = -0.58$, $SE = 1.29$, $t(82) = -0.45$, $p = .66$, $r_{\text{effect size}} = .05$, or perceived intelligence, $b = 1.42$, $SE = 1.29$, $t(82) = 1.10$, $p = .27$, $r_{\text{effect size}} = .12$.

Although the interaction between perceived intelligence and gender was not significant, we estimated the effect of perceived intelligence within each gender because of our a priori hypothesis. These exploratory analyses showed that perceived intelligence significantly predicted humanness for boys, $\beta = .29$, $t(82) = 2.47$, $p = .02$, but not girls, $\beta = .10$; $t(82) = 0.71$, $p = .48$.[2] The cross-classified model showed the same pattern: Perceived intelligence significantly predicted humanness for boys, $b = 2.10$, $SE = 0.85$, $t(82.10) = 2.48$, $p = .02$, $r_{\text{effect size}} = .26$, but not for girls, $b = 0.68$, $SE = 0.98$, $t(81.90) = 0.70$, $p = .49$, $r_{\text{effect size}} = .08$.

Discussion

Overall, these results provide evidence that attractiveness predicts children's humanness, and perceived intelligence predicts boys' humanness. Perceived intelligence did not significantly

[2] Although target age predicts children's humanness, $\beta = .20$, $t(81) = 3.37$, $p = .001$, including it in the model does not alter the magnitude or significance of the reported results.

predict boys' humanness more than girls' humanness, however. Because the effect sizes we observed here were smaller than those in Studies 1-3, and interaction effects require more power than main effects, the interaction effect between perceived intelligence and gender might have suffered from insufficient power (e.g., Shieh, 2008). We nevertheless observed some consistency with the adult data above, finding that attractiveness predicted ascriptions of both boys' and girls' humanness.

The results of Studies 1-3 suggest that people particularly consider women's attractiveness and men's intelligence when making very basic judgments of their humanity; the results of Study 4 suggest that some of these biased perceptions begin in childhood. Although these associations already entail alarming implications, they may have additional downstream consequences. We thus conducted a final study to examine whether the discrepant weight placed on men's and women's attractiveness and intelligence likewise predicts differences in moral treatment.

Study 5

Thus far, the studies show that people use both variations in targets' facial attractiveness and apparent intelligence to make inferences about their humanness. Further, these effects are gendered; participants consider more attractive-looking women and more intelligent-looking men to be more human than their unattractive and unintelligent-looking counterparts. The theoretical value of these results notwithstanding, understanding dehumanization carries important practical value by predicting meaningful life outcomes (see Kteily & Bruneau, 2017). The dehumanizing judgments we observed here could consequently speak to disparities in how society treats people. Namely, denying one's basic humanity often underlies moral mistreatment (e.g., Fincher & Tetlock, 2016; Gray et al., 2012). Given the gendered pattern of results, this may

render women especially vulnerable to moral mistreatment because of their unattractiveness and men especially vulnerable to moral mistreatment because of their perceived unintelligence. To test these possibilities, we investigated whether targets' gender, attractiveness, and perceived intelligence might bias participants' hypothetical decisions to sacrifice their lives.

To investigate moral decision-making, we used the classic "trolley problem"—a moral dilemma in which people decide whether to sacrifice the life of a single individual to save the lives of five others. To ensure variability in responses, we used the "Ned" version of the trolley problem in which approximately 55% of participants typically sacrifice the single individual (Hauser et al., 2008). We hypothesized that targets' perceived intelligence and attractiveness would negatively predict how many participants sacrifice them, and that attractiveness would more strongly predict decisions about women whereas perceived intelligence would more strongly predict decisions about men (https://osf.io/qxhzp/?view_only=55aab3cc523b4fef998959ef67126350).

Method

We recruited a new sample of 256 participants from MTurk to afford approximately 95% power for a cross-classified model (Westfall et al., 2014) with 412 targets and an assumed effect size of $r = .21$ (i.e., the average effect size in social and personality psychology; Richard et al., 2003). We randomly assigned the participants to rate either the male ($N = 125$; 77 female, 48 male; $M_{age} = 36.42$ years, $SD = 11.10$; 81 White, 13 Black, 12 Hispanic, 7 East Asian, 7 South Asian, and 5 "other") or female targets from Study 1 ($N = 107$; 61 female, 46 male; $M_{age} = 39.02$ years, $SD = 12.49$; 75 White, 15 Black, 7 East Asian, 5 Hispanic, 3 South Asian, and 2 "other") on the acceptability of sacrificing each to save the lives of others in the following scenario from Hauser et al. (2008, p. 285):

> A trolley is going down the tracks out of control. The driver of the trolley saw five people walking across the tracks and slammed on the brakes, but the brakes failed and the five people will not be able to get off the tracks in time. Fortunately, there is a switch which can temporarily turn the trolley onto a side track. There is a heavy object on the side track. If the trolley hits the object, the object will slow the trolley down and give the five people time to escape. Unfortunately, the heavy object is the individual pictured below, standing on the side track with his (her) back turned. If the switch is used, it will stop the trolley from killing the five people, but will kill the person pictured below. If the switch is not used, the five people will die. Is it acceptable to use the switch?

This description appeared above the face photo of each target, whom participants indicated how acceptable they believed it would be to sacrifice from 1 (*Not at all acceptable to use the switch*) to 7 (*Completely acceptable to use the switch*) in random order. Because participants rated many targets, we divided the trials into two blocks with a break between the two blocks. If fatigue during the study were to cause inattentive responding, we would expect an effect in the first block to weaken or become null in the second; we therefore tested whether the effects held over both blocks (if not, we would test whether they only hold in the first block).

We also had concerns that MTurk participants might be overly acquainted with the trolley problem and not respond naturally. To test this, we investigated whether an aggregate of participants' familiarity with the trolley problem (1 = *Not at all familiar*, 7 = *Extremely familiar*) and previous experience completing the trolley problem (1 = *Never*, 7 = *All of the time*) moderated our results, expecting the two measures to correlate.

Results

We first excluded data from 22 participants who either failed an attention check item by selecting "Decide between driving the trolley or having someone else drive the trolley" in response to the prompt "In the trolley situation, please indicate which description below best matches the decision you have to make" or who explicitly reported poor data quality (i.e., selected "I rated the faces randomly or I provided my ratings without seeing the faces" to two post-task probes), and two participants who skipped through the study.

Primary analyses. As predicted, trolley experience ($M = 1.96$, $SD = 1.32$) and trolley familiarity ($M = 4.19$, $SD = 2.36$) correlated, $r(230) = .49$, $p < .001$. We therefore aggregated these scores into a Trolley Knowledge composite. We then used a cross-classified model to predict the targets' acceptability of sacrifice from their gender; aggregated group-mean-centered attractiveness, perceived intelligence, and likeability scores from Study 1; their interactions and the participants' grand-mean-centered Trolley Knowledge.[3] Trolley Knowledge did not predict acceptability of sacrifice and is not discussed further.

Conceptually replicating Studies 1-4, we observed that less attractive and less intelligent-looking individuals were more willingly sacrificed (independent of their likeability; Table 12). Furthermore, women's attractiveness significantly predicted the acceptability of sacrificing them more than men's attractiveness predicted the acceptability of sacrificing them; likewise, men's perceived intelligence significantly predicted the acceptability of sacrificing them more than women's perceived intelligence predicted the acceptability of sacrificing them. Given that these

[3] For simplicity and efficiency, we do not report the analyses when interacting the predictors with participants' Trolley knowledge; all results remain substantively identical when doing so.

results emerged when including targets from both blocks, response fatigue is not a viable explanation for these data.

Table 12

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Targets' Acceptability of Sacrifice From Their Attractiveness, Perceived Intelligence, Likeability, Their Interactions, and Participants' Knowledge of the Trolley Problem

Predictors	*b*	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Target predictors						
Attractiveness	-0.06	0.01	408.97	-7.85	< .001	-.36
Perceived intelligence	-0.06	0.01	416.46	-4.70	< .001	-.22
Likeability	-0.07	0.01	420.52	-4.68	< .001	-.22
Gender	0.12	0.13	229.06	0.91	.36	.06
Attractiveness × Gender	0.04	0.01	405.04	6.06	< .001	.29
Perceived intelligence × Gender	-0.04	0.01	415.53	-3.04	.003	-.15
Participant predictors						
Trolley knowledge	-0.10	0.08	229	-1.22	.22	-.08

Note. Gender coded -1 = female, 1 = male.

Consistent with the studies above, we then estimated the model again using dummy coding to test the effects within each gender. As with the humanness ratings, attractiveness predicted women's acceptability of sacrifice but perceived intelligence did not. Perceived intelligence and attractiveness predicted men's acceptability of sacrifice (Table 13).

Table 13

Unstandardized Regression Coefficients, Standard Errors, and Test Statistics for the Model Predicting Male and Female Targets' Acceptability of Sacrifice From Their Attractiveness, Perceived Intelligence, Likeability, and Participants' Knowledge of the Trolley Problem

	b	*SE*	*df*	*t*	*p*	$r_{\text{effect size}}$
Female targets						
Attractiveness	-0.10	0.01	467.20	-9.23	< .001	-.39
Perceived intelligence	-0.02	0.02	467.20	-0.96	.34	-.04
Likeability	-0.08	0.02	467.20	-4.46	< .001	-.20
Trolley knowledge	-0.09	0.12	228	-0.73	.46	-.05
Male targets						
Attractiveness	-0.03	0.01	357.50	-2.31	.02	-.12
Perceived intelligence	-0.11	0.02	357.50	-6.10	< .001	-.31
Likeability	-0.05	0.02	357.50	-1.92	.06	-.10
Trolley knowledge	-0.11	0.11	228	-0.99	.32	-.07

Discussion

Here, targets' perceived intelligence and attractiveness predicted perceptions of the value of their lives. Consistent with the previous findings, men's and women's lives were valued differently, respective to their attractiveness and perceived intelligence (independent of their likeability). Put simply, people reported more willingness to sacrifice unattractive women and unintelligent-looking men to save the lives of five anonymous, unpictured strangers, quite literally devaluing their lives.

Although unattractive men also seemed more expendable, men's perceived intelligence predicted their dispensability three times better than their attractiveness did. Similarly, women's

attractiveness predicted their dispensability roughly seven times as much as their perceived intelligence did. Thus, parallel to the results observed above, attractiveness and intelligence particularly qualified perceptions of the human value of women's and men's lives, respectively.

Nevertheless, we did ask participants to make a hypothetical and extreme decision about others' life value over many trials (to ensure adequate variability in the targets' facial appearance). Future research should investigate this using a more realistic paradigm or by using archival data (e.g., criminal sentencing of men and women predicted by their facial appearance).

General Discussion

Recognizing people's humanity forms a vital basis for their humane treatment (e.g., Haslam, 2006). Emerging work demonstrates that bottom-up cues in people's facial appearance affect how human they seem. For instance, people rate faces that have more humanlike features (e.g., doll-like vs. human-like eyes), more humanlike structure (e.g., narrower and longer vs. wider and shorter), and that allow for configural processing (i.e., presented upright vs. inverted) as possessing more humanlike minds (see Deska & Hugenberg, 2017, for review). We advanced this emerging work by showing that higher-level trait impressions based on others' faces also affect how human they seem. In doing so, we help to answer the open question of why facial appearance affects how human people seem (Deska & Hugenberg, 2017), suggesting that one answer rests on whether they appear to possess traits that others value. We demonstrated this in two ways.

First, recognizing others' humanity requires recognizing them as intelligent agents (e.g., Gray et al., 2007; Haslam & Loughnan, 2014). Consistent with this, we found that less intelligent-looking men and women seem less human than their more intelligent-looking counterparts (Study 1). However, this effect was gendered. Men's perceived intelligence

predicted their humanness better than women's (Studies 1-3). In turn, the value placed on men's lives especially hinged on how intelligent they looked, such that people valued men's lives less to the extent that they looked unintelligent (Study 5). Looking unintelligent therefore renders one vulnerable to dehumanization, particularly for men. Future work should investigate the repercussions of biased attention towards men's intelligence when considering their humanity. For instance, the increasing numbers of men not entering (or finishing) university may face additional challenges resulting from seeming less human (e.g., Ferguson, 2016; Van Bavel et al., 2018). Likewise, men who have internalized gender stereotypes or who feel threatened by women's increasing professional successes might self-dehumanize when struggling academically or professionally (Willer et al., 2013).

We also found suggestive evidence that boys' humanness depends on their perceived intelligence. This may indicate that men's vulnerability to dehumanization based on their perceived intelligence begins early in life, though further work is needed to replicate this finding—particularly because this main effect did not significantly interact with the children's gender (i.e., although girls' perceived intelligence did not significantly predict their humanness, a nonsignificant interaction indicated no significant difference in the association between perceived intelligence and humanness in boys vs. girls). Should these findings replicate, they have troubling implications. For instance, boys who struggle academically or in ways that others would attribute to less self-control or sophisticated thinking (e.g., delinquency) may especially suffer dehumanization by others. The effect sizes for judgments of the young men in Studies 1-3 substantially exceeded those for the boys in Study 4, however, suggesting that trait-based dehumanization escalates as males mature. Thus, future work should examine whether boys encounter pronounced dehumanization for their perceived (un)intelligence as they feel increasing

pressure to act like adults during their transition to adulthood (e.g., to achieve financial independence; Arnett, 2001). Moreover, probing the developmental trajectory of these biases may reveal how early they develop (e.g., infancy, early childhood, preadolescence). Relatedly, investigating how these biases emerge in perceivers may also prove illuminating, such as by considering whether boys and girls are dehumanized differently during episodes of bullying (van Noorden et al., 2014).

Second, because society expects less attractive individuals to possess relatively undesirable traits (compared to attractive individuals), and because people highly value attractiveness in themselves and others (e.g., Dion et al., 1972), we expected less attractive individuals to seem less human. The results of all five studies supported this hypothesis. Given the breadth of the attractiveness halo effect, the association between people's attractiveness and perceptions of their humanness may not seem surprising. Yet, our findings quite clearly go beyond reiterating halo effects. First and most directly, all the effects we observed upheld when controlling for likeability. Second, research examining the attractiveness halo has typically described its effect on traits that vary between individuals, such as sociability (i.e., more attractive people seem more sociable; e.g., Dion et al., 1972). In distinction, although individuals' humanness (i.e., the extent to which they seem evolved) ought to apply equally to all people, individual differences in attractiveness still predicted judgments of how human they seemed. Finally, gender stereotypes exacerbated the effect for women (Studies 2 and 3). Thus, beyond reiterating the halo effect, these results suggest that (a) attractiveness guides people's impressions of the fundamental judgment of others' humanity (independent of likeability); (b) these judgments are especially biased in impressions of women; and (c) these judgments have

consequences for the perceived value of women's lives, such that people value the lives of unattractive women especially less than that of other individuals.

Because attractiveness negatively related to humanness judgments of both boys and girls in Study 3, our data suggest that women's acute vulnerability to their attractiveness in predicting their humanness may emerge between adolescence and early adulthood. However, we may have lacked sufficient power to fully detect a gender difference in the effect due to limits on the number of child targets we could acquire. Future work could thus help to clarify the developmental trajectory of this bias.

The prioritization of attractiveness over perceived intelligence in judgments of women's basic humanity is disturbing. The social gains that women have made through increased education and financial independence seem not to have realized in judgments of their basic humanity (e.g., Ferguson, 2016; Van Bavel et al., 2018). In addition to existing recommendations on how to prevent and reduce the objectification of women (e.g., understanding and altering the influence of media, informing the public, and conducting more research; American Psychological Association, 2007), more gains could be made by developing an intervention that equates women's attractiveness with their perceived intelligence in relation to their humanness. Beyond recapitulating gender stereotypes, our results show the dire need for such interventions. Otherwise, people may willingly and blatantly consider unattractive women and unintelligent-looking men as less than human.

Despite the consistent evidence observed here linking attractiveness to ascriptions of humanness, one could contend that attractiveness is not a particularly human trait because its physical basis renders it equally applicable to nonhuman entities. In some senses, we agree. Humanness constitutes more than a person's physique or face, and people's human value should

not be bound to their visual appeal. Yet our data clearly suggest people act otherwise. Participants conferred more humanness to attractive than to unattractive individuals in spite of the moral imperative to value all human lives. Attractiveness does not equally apply to human and nonhuman entities, however. Indeed, the processing of the physical attractiveness of human and nonhuman entities differs both qualitatively and quantitatively. The cues that make a tree, rose, or face visually appealing seem to be qualitatively different. Further, the social properties ascribed to a beautiful human would likely not adequately describe a beautiful frost-covered lake (e.g., extraverted, popular, robust). This distinction illuminates another important difference between human and nonhuman attractiveness: Whereas humans and nonhumans may share some physical precursors to an evaluation of high attractiveness (e.g., physical symmetry), other contributing factors do not seem as universal (e.g., a smooth skin complexion). Thus, both the antecedent physical and consequent social variables related to human attractiveness distinguish it from the attractiveness of non-human stimuli. Going further, even when attractive nonhuman entities and humans share some processes—such as symmetry making both more attractive—there are quantitatively different effects. For example, although symmetry appears to make various nonhuman targets more visually appealing, this effect is much stronger for human faces (e.g., Young et al., 2011). Thus, we believe that studying facial attractiveness is an essentially social psychological process, and one that shares more linguistic similarity than processing similarity with "attractive" nonhuman entities.

Second, although we agree that humanness should not be defined by physicality (people with many different bodies are all human), what appears true is that people do appear to use the human body as a cue to humanness. Physical traits are certainly understudied in the dehumanization literature; just as people have schemas about what human minds are, they also

have schemas about human appearance. Recent research indicates that perceivers dehumanize targets with different facial cues (e.g., high facial width-to-height ratio; Deska et al., 2018) or body types (e.g., short individuals; Kunst et al., 2019). Extending this, we find here that people have a prototype for what humans should look like based on physical attractiveness and that deviations from this prototype are met with dehumanization (albeit more so for women than for men in our data). A range of other physical traits (untested in the current work but exciting for future research) might also trigger dehumanizing judgments from this perspective, including physical disfiguration, obesity, mobility impairment, and others. Thus, a key strength of the current work is in connecting previously distinct research areas to the broader conceptualization of humanness.

Limitations and Future Directions

Despite the consistent results observed in this work, limitations remain which may provide fruitful directions for future research. First, we focused on judgments of White individuals to avoid the potential influence of race-based dehumanization among our primarily White participant population (e.g., Goff et al., 2008). Expanding this investigation to examine how race intersects with our observations is thus an important future step. For instance, because race-based stereotypes particularly differ regarding competence (Fiske et al., 2002), members of racial groups stereotyped as competent may especially suffer dehumanization when they fail to live up to their group's stereotypes, much as unintelligent-looking men did in our studies.

Second, besides focusing on gender, the current work limited its investigation to the traits of attractiveness and intelligence based on perceptions of individuals' faces. Although these traits emerged as potential triggers of dehumanization based on the extant literature, other face-

based trait inferences likely affect ascriptions of humanness as well (e.g., dominance, trustworthiness; Wilson et al., 2018).

Third, although we used a well-validated scale to measure humanness (and further validated it here by distinguishing the key results from likeability; see Kteily et al., 2015), future work should develop another scale immune to the limitations associated with existing measures of humanness (e.g., conflating the dependent and independent variables), to further test the generality of our findings.

Fourth, the present work focuses only on judgments of static faces. Judgments of traits arise from other cues too, however, such as body movements, vocal tone, and dynamic facial cues (e.g., Schroeder & Epley, 2015; Thoresen et al., 2012). Thus, future work could consider whether attractiveness, intelligence, and potential other traits lend themselves to judgments of humanness across other modalities as well, helping to illuminate the role of dynamic nonverbal cues in ascriptions of humanness more broadly.

Fifth, our adult targets had a relatively restricted age range, whereas the child targets in Study 4 had a relatively wide range from a developmental perspective (i.e., 3-12 years old). Certain traits can hold specific value across the lifespan, therefore predicting individuals' humanness differently depending on their age (e.g., agreeableness rises in value as one ages; Hudson & Fraley, 2016).

Sixth, although we consistently found gendered effects of attractiveness and intelligence, some of the results in Studies 2 and 3 which were designed to understand these gendered effects were not consistent across the studies. However, meta-analyzing the results across the two studies showed greater consistency. For female targets, valuing women's attractiveness predicted the strength of the attractiveness-humanness relationship. For male targets, valuing men's

intelligence predicted the strength of the intelligence-humanness relationship. Thus, gendered stereotypes appear to be at play in this gendered dehumanization. Surprisingly, though, this meta-analysis also showed that valuing men's intelligence predicted this attractiveness-humanness relationship. Although speculative, this may be due in part to the way in which people infer cues from faces. For male faces, attractiveness signals competence, meaning a masculine face can cue both dimensions. However, for female faces, attractiveness is strongly positively correlated with femininity, which itself is negatively correlated with inferences of competence (paralleling research on role congruity theory and the "double bind;" e.g., Diekman et al., 2010; Eagly & Karau, 2002). Thus, valuing men's intelligence may actually involve valuing overlapping facial signals of attractiveness and intelligence.

Finally, we did not examine whether face-based humanness ascriptions influence actual social interactions. We observed evidence in Study 5 that judgments of attractiveness and intelligence have downstream consequences for evaluating individuals' moral value, albeit in a situation with limited external validity. Future work should thus investigate whether humanness judgments affect actual behavior, especially in contexts that avail information beyond facial appearance.

Conclusions

The current work suggests that individuals' facial appearance affects how human they seem. Unintelligent-looking men and unattractive women are especially vulnerable to being dehumanized and, subsequently, their lives are systematically valued less. Importantly, these effects occur independent of targets' likeability. These data provide greater nuance to the mechanisms underlying dehumanization and may help to explain how intergroup biases support gender-based dehumanization.

References

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

American Psychological Association, Task Force on the Sexualization of Girls. (2007*). Report of the APA Task Force on the Sexualization of Girls*. Retrieved from http://www.apa.org/pi/wpo/sexualization.html.

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play!. *Science*, *323*, 1183–1183.

Arnett, J. J. (2001). Conceptions of the transition to adulthood: Perspectives from adolescence through midlife. *Journal of Adult Development*, *8*, 133–143.

Badawy, R. L., Gazdag, B. A., Bentley, J. R., & Brouer, R. L. (2018). Are all impostors created equal? Exploring gender differences in the impostor phenomenon-performance link. *Personality and Individual Differences*, *131*, 156–163.

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion*. Chicago, IL: Rand McNally.

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex- role stereotypes: A current appraisal. *Journal of Social Issues*, *28*, 59–78.

Bruneau, E., Jacoby, N., Kteily, N., & Saxe, R. (2018). Denying humanity: The distinct neural correlates of blatant dehumanization. *Journal of Experimental Psychology: General*, *147*, 1078–1093.

Buss, D. M. (1988). The evolution of human intrasexual competition: Tactics of mate attraction. *Journal of Personality and Social Psychology*, *54*, 616–628.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*, 1–14.

Buss, D. M. (1994). *The evolution of desire*. New York, NY: Harper Collins.

Buss, D. M., & Dedden, L. A. (1990). Derogation of competitors. *Journal of Social and Personal Relationships*, *7*, 395–422.

Butler, D., & Geis, F. L. (1990). Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social Psychology*, *58*, 48–59.

Capozza, D., Andrighetto, L., Di Bernardo, G. A., & Falvo, R. (2012). Does status affect intergroup perceptions of humanity?. *Group Processes & Intergroup Relations*, *15*, 363–377.

Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the trolley track: Neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience*, *5*, 404-413.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., ... & Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*, 1355–1360.

Deska, J. C., Almaraz, S. M., & Hugenberg, K. (2017). Of mannequins and men: Judgments of animacy in faces are bounded by perceptual and processing similarities to human faces. *Social Psychological & Personality Science*, *8*, 183–190.

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018). Facing humanness: Facial width-to-height ratio predicts ascriptions of humanity. *Journal of Personality and Social Psychology*, *114*, 75–94.

Deska, J. C. & Hugenberg, K. (2017). The face-mind link: Why we see minds behind faces, and how others' minds change how we see their face. *Social and Personality Psychology Compass*, *11*, e12361.

Diekman, A. B. (2007). Negotiating the double bind: Interersonal and instrumental evaluations of dominance. *Sex Roles, 22,* 551–561.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science, 21,* 1051–1057.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*, 285–290.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109–128.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*, 573–598.

Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research*, *20*, 145–156.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864–886.

Farina, A., Fischer, E. H., Sherman, S., Smith, W. T., Groh, T., & Mermin, P. (1977). Physical attractiveness and mental illness. *Journal of Abnormal Psychology*, *86*, 510–517.

Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attraction: A comparison across five research paradigms. *Journal of Personality and Social Psychology*, *59*, 981–993.

Ferguson, S. J. (2016). Women and Education: Qualifications, skills, and technology (Research Report No. 89-503-X). Retrieved from Statistics Canada website: https://www150.statcan.gc.ca/n1/pub/89-503-x/2015001/article/14640-eng.htm

Fincher, K. M., Kteily, N. S., & Bruneau, E. G. (2018). Our humanity contains multitudes: Dehumanization is more than overlooking mental capacities. *Proceedings of the National Academy of Sciences*, *115*, E3329-E3330.

Fincher, K. M., & Tetlock, P. E. (2016). Perceptual dehumanization of faces is activated by norm violations and facilitates norm enforcement. *Journal of Experimental Psychology: General*, *145*, 131–146.

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, *11*, 154–158.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902.

Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly*, *21*, 173–206.

Furnham, A., & Rawles, R. (1995). Sex differences in the estimation of intelligence. *Journal of Social Behavior and Personality*, *10*, 741–748.

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*, *15*, 380-387.

Gillen, M. M., & Lefkowitz, E. S. (2009). Emerging adults' perceptions of messages about physical appearance. *Body Image*, *6*, 178–185.

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology*, *94*, 292–306.

Goff, P. A., Jackson, M. C., Leone, D., Lewis, B. A., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing Black children. *Journal of Personality and Social Psychology*, *106*, 526–545.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*, 619–619.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101–124.

Gunaydin, G., Selcuk, E., & Zayas, V. (2017). Impressions based on a portrait predict, 1-month later, impressions following a live interaction. *Social Psychological and Personality Science*, *8*, 36–44.

Gwinn, J. D., Judd, C. M., & Park, B. (2013). Less power = less human? Effects of power differentials on dehumanization. *Journal of Experimental Social Psychology*, *49*, 464–470.

Hagan, R. L., & Kahn, A. (1975). Discrimination against competent women. *Journal of Applied Social Psychology*, *5*, 362–376.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science*, *17*, 847–853.

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*, 45–51.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252–264.

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*, 399–423.

Hauser, M. D., Young, L., & Cushman, F. A. (2008). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral psychology and biology*. New York, NY: Oxford University Press.

Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology*, *113*, 513–529.

Hehman, E., Xie, S. Y., Ofosu, E. K., & Nespoli, G. A. (2018, February 19). Assessing the point at which averages are stable: A tool illustrated in the context of person perception. http://doi.org/10.17605/OSF.IO/2N6JQ.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*, 657– 674.

Heilman, M. E., & Saruwatari, L. R. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs. *Organizational Behavior and Human Performance*, *23*, 360–372.

Heilman, M. E., & Stopeck, M. H. (1985). Being attractive, advantage or disadvantage? Performance-based evaluations and recommended personnel actions as a function of appearance, sex, and job type. *Organizational Behavior and Human Decision Processes*, *35*, 202–215.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185.

Horner, M. S. (1972). Toward an understanding of achievement related to conflicts in women. *Journal of Social Issues*, *28*, 157–175.

Hudson, N. W., & Fraley, R. C. (2016). Do people's desires to change their personality traits vary with age? An examination of trait change goals across adulthood. *Social Psychological and Personality Science*, *7*, 847–856.

Hugenberg, K., Young, S., Rydell, R. J., Almaraz, S., Stanko, K. A., See, P. E., & Wilson, J. P. (2016). The face of humanity: Configural face processing influences ascriptions of humanness. *Social Psychological and Personality Science*, *7*, 167–175.

Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*, 719–737.

Jackson, L. A. (1992). Physical appearance and gender: Sociobiological and sociocultural perspectives. Albany, NY: State University of New York Press.

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, *83*, 518–527.

Kalof, L. (1999). Stereotyped evaluative judgments and female attractiveness. *Gender Issues*, *17*, 68–82.

Khalid, S., Deska, J. C., & Hugenberg, K. (2016). The eyes are the windows to the mind: Direct eye gaze triggers the ascription of others' minds. *Personality and Social Psychology Bulletin*, *42*, 1666–1677.

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*, 543–555.

Krendl, A. C., Macrae, C. N., Kelley, W. M., Fugelsang, J. A., & Heatherton, T. F. (2006). The good, the bad, and the ugly: An fMRI investigation of the functional anatomic correlates of stigma. *Social Neuroscience*, *1*, 5–15.

Krumhuber, E., Lai, Y., Rosin, P. L., & Hugenberg, K. (2018). When facial expressions do and do not signal minds: The role of face inversion, expression dynamism, and emotion type. *Emotion*. Advanced online publication.

Kteily, N., & Bruneau, E. (2017). Backlash: The politics and real-world consequences of minority group dehumanization. *Personality and Social Psychology Bulletin*, *43*, 87–104.

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, *109*, 901–931.

Kunst, J. R., Kteily, N., & Thomsen, L. (2019). "You little creep": Evidence of blatant dehumanization of short groups. *Social Psychological and Personality Science*, *10*, 160–171.

Neter, J., Wasserman, W., & Kutner, M. H. (1990). *Applied linear statistical models* (3rd ed.). Chicago, IL: Irwin.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*, 1377–1388.

Lerner, R. M., Orlos, J. B., & Knapp, J. R. (1976). Physical attractiveness, physical effectiveness and self-concept in late adolescents. *Adolescence*, *11*, 313–326.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*, 262–265.

Meyer, M., Cimpian, A., & Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, *6*, 235.

Leyens, J. P., Demoulin, S., Vaes, J., Gaunt, R., & Paladino, M. P. (2007). Infra-humanization: The wall of group differences. *Social Issues and Policy Review*, *1*, 139–172.

Leyens, J. P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M. P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, *31*, 395–411.

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality & Social Psychology*, *82*, 947–955.

LoBue, V. & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) Set: Validity and reliability from untrained adults. *Frontiers in Emotion Science, 5.*

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy. How, when, and where we perceive life in a face. *Psychological Science*, *21*, 1854 –1862.

Meyer, M., Cimpian, A., & Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, *6*, 235.

Miceli, M., & Castelfranchi, C. (2018). Contempt and disgust: The emotions of disrespect. *Journal for the Theory of Social Behaviour*, 1–25.

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, *5*, 498–502.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*, 867-872.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*, 1–20.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D., Abdi, H. (1998). The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition*, *26*, 146–160.

Pardoe, I. (2000). An introduction to bootstrap methods using Arc. Retrieved May 3, 2019, from the University of Minnesota, School of Statistics Web site: (https://www.stat.umn.edu/arc/bootmeth.pdf).

Pereira, C., Vala, J., & Leyens, J. P. (2009). From infra-humanization to discrimination: The mediation of symbolic threat needs egalitarian norms. *Journal of Experimental Social Psychology*, *45*, 336–344.

Porter, N., & Geis, F. L. (1981). Women and nonverbal leadership cues: When seeing is not believing. In C. Mayo and N. Henley (Eds.), *Gender and nonverbal behavior* (pp. 39–61). New York: Springer Verlag.

Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T., & Heatherton, T. F. (2014). Social connection modulates perceptions of animacy. *Psychological Science*, *25*, 1943–1948.

Praino, R., Stockemer, D., & Ratis, J. (2014). Looking good or looking competent? Physical appearance and electoral success in the 2008 congressional elections. *American Politics Research*, *42*, 1096–1117.

Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. *Cognition and Emotion*, *25*, 140–148.

Prokosch, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate choice: Intelligent men are always appealing. *Evolution and Human Behavior*, *30*, 11–20.

Rai, T. S., Valdesolo, P., & Graham, J. (2017). Dehumanization increases instrumental violence, but not moral violence. *Proceedings of the National Academy of Sciences*, *114*, 8511-8516.

Raty, H. and Snellman, L., 1992. Does gender make any difference? Common sense conceptions of intelligence. *Social Behavior and Personality*, *20*, 23–34.

Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex- typicality and attractiveness: Are supermale and superfemale faces super- attractive?. *British Journal of Psychology*, *91*, 125–140.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, *3*, 16–20.

Rubenstein, A. J., Kalakanis, L., & Langlois, J. H. (1999). Infant preferences for attractive faces: A cognitive explanation. *Developmental Psychology*, *35*, 848–855.

Rule, N. O., Bjornsdottir, R. T., Tskhay, K. O., & Ambady, N. (2016). Subtle perceptions of male sexual orientation influence occupational opportunities. *Journal of Applied Psychology*, *101*, 1687–1704.

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, 104, 409–429.

Schein, S. S., & Langlois, J. H. (2015). Unattractive infant faces elicit negative affect from adults. *Infant Behavior and Development*, *38*, 130–134.

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, *26*, 877–891.

Sherman, G. D., & Haidt, J. (2011). Cuteness and disgust: The humanizing and dehumanizing effects of emotion. *Emotion Review*, *3*, 245–251.

Shieh, G. (2008). Detecting interaction effects in moderated multiple regression with continuous variables. *Organizational Research Methods*, *12*, 1–19.

Signorielli, N. (1990). Children, television, and gender roles: Messages and impact. *Journal of Adolescent Health*, *11*, 50–58.

Smiler, A. P., & Epstein, M. (2010). Measuring gender: Options and issues. In J. C. Chrisler, & D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 133–157). New York: Springer.

Smith, M., & Wilhelm, J. D. (2004). "I just like being good at it": The importance of competence in the literate lives of young men. *Journal of Adolescent & Adult Literacy*, *47*, 454–461.

Striegel-Moore, R. H., Silberstein, L. R., & Rodin, J. (1986). Toward an understanding of risk factors for bulimia. *American Psychologist*, *41*, 246–263.

Sui, J., & Liu, C. H. (2009). Can beauty be ignored? Effects of facial attractiveness on covert attention. *Psychonomic Bulletin & Review*, *16*, 276–281.

Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. Cognition, *124*, 261–271.

Tiggemann, M., & Lewis, C. (2004). Attitudes toward women's body hair: Relationship with disgust sensitivity. *Psychology of Women Quarterly*, *28*, 381–387.

Tiggemann, M., & Lynch, J. E. (2001). Body image across the life span in adult women: The role of self-objectification. *Developmental Psychology*, *37*, 243–253.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*, 813–833.

Unger, R. K. (1979). *Female and male*. New York, NY: Harper and Row.

Van Bavel, J., Schwartz, C., & Esteve, A. (2018). The reversal of the gender gap in education and its consequences for family life. *Annual Review of Sociology*, *44*.

van Noorden, T. H., Haselager, G. J., Cillessen, A. H., & Bukowski, W. M. (2014). Dehumanization in children: The link with moral disengagement in bullying and victimization. *Aggressive Behavior*, *40*, 320–328.

Viki, G. T., Osgood, D., & Phillips, S. (2013). Dehumanization and self-reported proclivity to torture prisoners of war. *Journal of Experimental Social Psychology*, *49*, 325–328.

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*, 383–388.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*, 2020–2045.

Willer, R., Rogalin, C. L., Conlon, B., & Wojnowicz, M. T. (2013). Overdoing gender: A test of the masculine overcompensation thesis. *American Journal of Sociology*, *118*, 980–1022.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal sentencing outcomes. *Psychological Science*, *26*, 1325–1331.

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

Wilson, J. P., Young, S. G., Rule, N. O., & Hugenberg, K. (2018). Configural processing and social judgments: Face inversion particularly disrupts inferences of human-relevant traits. *Journal of Experimental Social Psychology, 74*, 1–7.

Young, S. G., Sacco, D. F., & Hugenberg, K. (2011). Vulnerability to disease is associated with a domain- specific preference for symmetrical faces relative to symmetrical non- face stimuli. *European Journal of Social Psychology*, *41*, 558–563.

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, 26, 237–242.

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, *28*, 238–249.

Appendix

1. In order for a man (woman) to be trusted, I think it is important that he (she) is intelligent.
2. In order for a man (woman) to be liked, I think it is important that he (she) is intelligent.
3. In order for a man (woman) to get along with others, I think it is important that he (she) is not intelligent (reverse-scored).
4. In order for a man (woman) to be trusted, I think it is important that he (she) is physically attractive.
5. In order for a man (woman) to be liked, I think it is important that he (she) is physically attractive.
6. In order for a man (woman) to get along with others, I think it is important that he (she) is not physically attractive (reverse-scored).
7. In order for a man (woman) to become wealthy, I think it is important that he (she) is intelligent.
8. In order for a man (woman) to become successful in his (her) career, I think it is important that he (she) is intelligent.
9. In order for a man (woman) to get ahead in his (her) career, I think it is important that he (she) is not intelligent (reverse-scored).
10. In order for a man (woman) to become wealthy, I think it is important that he (she) is physically attractive.
11. In order for a man (woman) to become successful in his (her) career, I think it is important that he (she) is physically attractive.
12. In order for a man (woman) to get ahead in his (her) career, I think it is important that he (she) is not physically attractive (reverse-scored).

13. Generally speaking, I think it is more important for a man (woman) to be intelligent than physically attractive.
14. Generally speaking, I think it is more important for a man (woman) to be physically attractive than intelligent.