SUPPLEMENTARY METHODS

Study 1

Method

Open Science

We report all measures, manipulations, and exclusions we have taken, and all data, analysis code, experimental materials, and supplementary results are available for download at: https://osf.io/sehrn. Our first study was primarily exploratory, and therefore was not pre-registered.

Ethics Statement

Relevant ethical guidelines were followed and all studies in this paper were approved through University of XX's Central University Research Ethics Committee, with the reference number XXXX.

Participants

We recruited 301 participants based in the U.S. via Amazon Mechanical Turk (MTurk). Data collection occurred in November 2017, and participants were paid $1.20 for their time (determined through the survey taking approximately 10 minutes, with an hourly US minimum wage of $7.25). Participants could not take part in the survey if they had participated in related studies by us in the past, and were excluded from analysis if they did not complete the survey in full (N=2), took the survey more than once (N=3) or failed a simple attention check asking them to indicate the beliefs of other player: the 'target' (N=19). This left us a final sample of 275 participants.

Our sample size of 300 was determined by a power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009), which indicated we would need at least 274 participants

to detect a small-to-medium effect of f=0.17, taking an a of .05 and power of .80 (with f=0.10 being conventionally small, and f=0.25 being conventionally medium).

The majority of participants identified as White (n=207), followed by Hispanic (n=22), Asian (n=21), and Black (n=21), and on a scale of 1 (very liberal) to 7 (very conservative), the mean ideology score was 3.19 (SD=1.60). Overall, participants scored fairly low on both speciesism (M=3.06, SD=1.38; scale α=.89) and racism (M=2.67, SD=1.52; scale α=.91), and as in previous research these were significantly positively correlated (r=0.34, $p < .001$). All participants were included in data analysis, regardless of ethnicity. Unfortunately, due to a coding error we did not collect data on participant's age, gender, or political affiliation.

Design

This study had a 2 (Prejudice Type: Speciesism vs. Racism) x 2 (Target: Prejudiced vs. Non-Prejudiced) between-subjects design, where participants were asked to rate a target who expressed either racist, speciesist, non-racist or non-speciesist beliefs.

Procedure

At the start of the study, participants completed a measure of their own prejudicial attitudes. Participants in the speciesism conditions completed the six-item (α=.89) speciesism scale (Caviola et al., 2018), and participants in the racism conditions completed a condensed[1] four-item version (α=.91) of the modern racism scale (McConahay, 1986). Participants were then told that in a previous survey, we asked other MTurk participants to complete these same scale questions and that "In this study, we will tell you how one such person responded to a

[1] Note that the original version consists of seven items, but we selected only four of these. While some items are still highly relevant to popular discourse in 2018 (e.g. "Blacks are getting too demanding in their push for equal rights"), we were concerned that others appeared anachronistic (e.g. "Blacks have more influence upon school desegregation plans than they ought to have") or too explicitly racist in a modern context (e.g. "Over the past few years, the government and news media have shown more respect to blacks than they deserve"). Given this, we selected a subset of only four items. The four items still showed high internal reliability (α=.91).

few of these previous questions about [black people / animals / men and women]. You will be answering some questions about your impressions of this person based on their answers."

Participants were always presented with three questions from the full list they had answered, where the target either agreed or disagreed with these items drawn from the scales. Participants in the non-speciesist condition were told that the target had strongly disagreed (scale point 1) that "Morally, animals always count for less than humans"; had disagreed (scale point 2) that "Humans have the right to use animals however they want to"; and strongly agreed (scale point 7; item is reverse coded) that "Chimpanzees should have basic legal rights such as a right to life or a prohibition of torture". Participants in the pro-speciesism condition were given the same three statements, but the level of agreement was reversed: the speciesist strongly agreed to the first statement (7 instead of 1), agreed to the second (6 instead of 2), and strongly disagreed with the third (1 instead of 7). Participants in the pro-racism condition were told that the target had strongly agreed (scale point 7) that "Discrimination against black people is no longer a problem in the United States"; had agreed (scale point 6) that "Black people are getting too demanding in their push for equal rights"; and strongly disagreed (scale point 1; item reverse coded) that "It is easy to understand the anger of black people in America". Again, participants in the anti-racism condition were given the same three statements, but the level of agreement was reversed. To ensure that participants had read and understood this information, they were given two attention check questions in which they were required to report whether the target agreed or disagreed with the first two statements (e.g. "*Did the other person agree or disagree that 'Morally, animals always count for less than humans"*?"). Participants who answered either of these questions incorrectly were excluded from analysis.

Measures

First, participants completed a number of character ratings, rating the target in terms of how moral (*1 = extremely immoral / bad, 7 = extremely moral / good*), trustworthy (*1 =*

extremely untrustworthy, 7 = extremely trustworthy), kind (*1 = extremely unkind, 7 = extremely kind*), warm (*1 = extremely cold, 7 = extremely warm*), sociable (*1 = not at all sociable, 7 = extremely sociable*) competent (*1 = not at all competent, 7 = extremely competent*), capable (*1 = not at all capable, 7 = extremely capable*), loyal (*1 = extremely disloyal, 7 = extremely loyal*), reliable (*1 = extremely unreliable, 7 = extremely reliable*), radical (*1= extremely un-radical, 7 = extremely radical*), and judgmental (*1 = not at all judgmental, 7 = extremely judgmental*) they were expected to be. The four items assessing how moral the target was thought to be (moral, kind, trustworthy, loyal) were combined into a single measure of perceived morality (α=.93), as were the two competence items (competent, capable: α=.89) and the two warmth items (warm, sociable: α=.82).

Second, participants completed some role suitability questions, rating on a 1-7 scale how suitable the target would be for six social roles: friend, romantic partner, work colleague, boss, CEO, and political leader (e.g. *1=an extremely bad friend, 7=an extremely good friend*). These six items showed high internal consistency, and so were combined into a single measure of perceived suitability as a social partner (α=.96). Analyses with each item individually can be seen in the Supplementary Information.

Third, participants answered two questions about the ideology and belief of the target, rating how politically liberal or conservative they thought the target to be (*1=very liberal / left, 7=very conservative / right),* and how religious they thought the target (*1=not at all religious, 7=very religious*).

Fourth, participants rated on a 1-7 scale how much the target would support three progressive "liberation", or "rights" movements (gay rights, black rights, animal rights). We were primarily interested in the ratings of perceived support for liberation movements for the non-manipulated group: in all conditions for participants perceptions of the target's support for

gay rights, and then for participants in the speciesism conditions perceived support for black rights, and for participants in the racism conditions perceived support for animal rights.

Study 2

Method

Open Science

As for all studies in this paper, report all measures, manipulations, and exclusions, and all data, analysis code, experimental materials, and supplementary results are available for download at: https://osf.io/5wscp. Our design, analysis plan, and hypotheses were pre-registered on the Open Science Framework. The pre-registration can be seen at: https://osf.io/fyu6s.

Participants

We had 451 participants based in the U.S. complete the survey via MTurk. Data collection occurred in December 2017, and participants were paid $1.20 for their time. Participants could not take part in the survey if they had participated in related studies by us in the past, and in accordance with the pre-registration were excluded from analysis if they did not complete the survey in full (N=6), took the survey more than once (N=13) or failed a simple attention check asking them to indicate the beliefs of the target (N=22). This left us a final sample of 410 participants, which was sufficient to detect a small-to-medium size effect (f=0.16), taking an a of .05 and power of .80[2].

[2] In the course of writing up the manuscript we realized that the a priori power analysis as reported in the pre-registration for Study 2 contained a small error. Our pre-registered power analysis indicated we would need 403 participants to detect a small-to-medium size effect of f=0.14 (with f=0.10 being conventionally small, and f=0.25 being conventionally medium), taking an a of .05 and power of .80. However, in the course of doing the power analysis for the third study, we realised that this power analysis was incorrect because it did not account for the covariate (participant's own prejudice) included in the analyses, and the numerator *df* had been incorrectly specified: with three prejudice types, the numerator *df* should have been 2, but we originally entered 1 by accident. All this meant that in actuality we had power to detect a small-to-medium effect size of f=0.16, not f=0.14 – which still relates to -a small-to-medium effect size, so we deem this change as minor.

The majority of participants identified as White (*n*=297), followed by Black (*n*=39), Hispanic (*n*=36), and Asian (*n*=21). We had roughly equal numbers of female (*n*=197) and non-female participants (*n*=213), and the mean age was 41 years old. On average, participants were politically left-to-moderate (*M*=3.30, *SD*=1.75), with more participants identifying as Democrat (*n*=182) or Independent (*n*=125) than Republican (*n*=80). Participants scored fairly low on racism (*M*=2.71, *SD*=1.43; scale α=.93), sexism (*M*=3.17, *SD*=1.45; scale α=.94), and speciesism (*M*=3.16, *SD*=1.32; scale α=87). As in previous research, participants' speciesism was positively correlated with their racism (r=0.39, $p < .001$) and sexism (r=0.37, $p < .001$), and these were in turn positively correlated with each other (r=0.69, $p < .001$). All participants were included in analysis regardless of their demographic features.

Design

This study had a 3 (Prejudice Type: Speciesism vs. Racism vs. Sexism) x 2 (Target: Prejudiced vs. Non-Prejudiced) between-subjects design. The instructions and items used for the racism and speciesism conditions were identical to those used in Study 1. For participants in the sexism conditions, we used statements taken from the hostile sexism[3] subscale of the Ambivalent Sexism Inventory (Glick & Fiske, 1996). Participants in the sexism condition were told that the other target had strongly agreed (scale point 7) that "Women are too easily offended"; had agreed (scale point 6) that "When women lose to men in a fair competition, they typically complain about being discriminated against"; and strongly disagreed (scale point 1; item reverse coded) that "Feminists are making entirely reasonable demands of men". Again, participants in the anti-sexism condition were given the same three statements, but the level of

[3] Items from the hostile sexism subscale were chosen instead of items from the benevolent sexism subscale, as it is the hostile sexism items that are most comparable to the 'typical', obviously negative prejudice as measured by the modern racism and speciesism scales. As applied to racial prejudice, 'benevolent' racial prejudice might include the belief that black people have special qualities (e.g. musically talented or athletic) that seem positive on first glance but actually enforce a hierarchy which places them in a subservient position. As applied to speciesism, 'benevolent' speciesist prejudice might include the belief that animals need protecting and that while 'cute' do not have worth outside of their status as property of people.

agreement was reversed so that the non-sexist strongly disagreed with the first statement (1 instead of 7), disagreed to the second (2 instead of 6), and strongly agreed with the third (7 instead of 1).

Measures

The measures used in Study 2 were almost identical to those in Study 1, with a few exceptions. First, in the interests of time we removed four of the least relevant character ratings from Study 1 (radical, reliable, trustworthy, loyal), which seemed most redundant to other items included. Second, we added two new questions where we asked participants to a) guess the target's gender, and b) predict the target's social dominance. To measure predicted social dominance orientation (SDO), we asked participants "How much do you think this person would agree or disagree with the statement that "It is probably a good thing that certain groups are at the top and other groups are at the bottom"? (*1=they would strongly disagree, 7=they would strongly agree*). To measure predicted gender, we asked participants whether they thought it more likely the target was male, more likely to be female, or equally likely to be male or female.

Study 3

Method

Open Science

As for all studies in this paper, report all measures, manipulations, and exclusions, and all data, analysis code, experimental materials, and supplementary results are available for download at: https://osf.io/fvux3. Our design, analysis plan, and hypotheses were pre-registered on the Open Science Framework, and the pre-registration can be seen at: https://osf.io/dp98h.

Participants

We had 432 participants based in the U.S. complete the survey via MTurk. Data collection occurred in January 2018, and participants were paid $1.20 for their time. Participants could not take part in the survey if they had participated in related studies by us in the past, and in accordance with the pre-registration were excluded from analysis if they did not complete the survey in full (N=1), took the survey more than once (N=8) or failed a simple attention check asking them to indicate the beliefs of the target (N=20). This left us a final sample of 403 participants.

As outlined in our pre-registration, we actually wanted to recruit 340 participants to detect a small-to-medium size effect (f=0.16), taking an a of .05 and power of .80. Unfortunately, a typo in the set-up on MTurk meant we recruited 430 participants instead of 340 participants. Note that we did not conduct any analyses until the full data collection was complete, and this error in recruiting a larger sample than planned actually gave us more power to detect effects.

The majority of participants identified as White (n=306), followed by Black (n=37), Asian (n=29), and Hispanic (n=20). We had roughly equal numbers of female (n=194) and non-female participants (n=209), and the mean age was 41 years old. On average, participants were politically left-to-moderate (M=3.38, SD=1.70), with more participants identifying as Democrat (n=178) or Independent (n=118) than Republican (n=78). Participants scored fairly low on homophobia (M=2.69, SD=1.88; scale α=.92) and speciesism (M=3.02, SD=1.33; scale α=.87), and these were significantly positively correlated (r=.28, $p < .001$). All participants were included in analysis regardless of their demographic features.

Design

This study had a 2 (Prejudice Type: Speciesism vs. Homophobia) x 2 (Target: Prejudiced vs. Non-Prejudiced) between-subjects design. The instructions and items used for

the speciesism conditions were identical to those used in Studies 1 and 2. For participants in the homophobia conditions, we used statements taken from the 5-item Attitude Towards Gay Men scale (Herek, 1998). Participants in the pro-homophobia condition were told that the other target had strongly agreed (scale point 7) that "Male homosexuality is a perversion[4]"; had agreed (scale point 6) that "Sex between two men is just plain wrong"; and strongly disagreed (scale point 1; item reverse scored) that "Male homosexuality is merely a different kind of lifestyle that should not be condemned" Again, participants in the anti-homophobia condition were given the same three statements, but the level of agreement was reversed so that the non-homophobe strongly disagreed with the first statement (1 instead of 7), disagreed to the second (2 instead of 6), and strongly agreed with the third (7 instead of 1).

Measures

The measures used in Study 3 were almost identical to those used in Study 2. While we again measured perceptions of warmth, competence, and morality, given that our previous results were robust across the individual items, in the interests of space we only used a single item for each (moral; warm or cold; competent). Similarly, we again measured suitability for different social roles, but given that our previous results were robust across the individual items, for reasons of space used only four instead of six roles (suitability as a friend; romantic partner; boss; political leader: scale a=.94).

In the DG, participants were told that they had an additional bonus of $0.30 and that they could choose to transfer some of this amount to the target, which would be paid to them as a bonus after the study. Choices were given in 5 cent increments, and participants were given the amounts that each target would receive in parentheses (e.g. "*Give 0 (You 30, Other 0)";*

[4] It was noted by a reviewer that this first item – "Male homosexuality is a perversion"- is quite extreme. However, participants own scores for this item (*M*=2.61, *SD*=2.04) were comparable to the other items, and this item actually did not yield the lowest average score. Moreover, when looking only at Republican participants, we observed higher scores around the mid-point for this item (*M*=4.22, *SD*=2.09), like for the others. It seems that participants themselves did not perceive this as a particularly extreme item (see supplementary results for more information).

"Give 5 (You 25, Other 5)"). This enabled us to get a behavioural measure of participants behavioural intentions towards the target. At the end of the study, participants were actually paid based on their decision, receiving the $0.30 bonus minus whatever they had chosen to transfer.