**Study 1**

**Method**

I recruited 125 people from Amazon’s Mechanical Turk (MTurk) to participate in a “survey on perceiving others’ attitudes.” This sample size allows 80% power to detect an effect of *r* = .25 and 90% power to detect an effect of *r* = .30. These correlations were informed by being on the lower-bound of relevant past research (White & Crandall, 2017). A total of 126 people participated. Participants’ ages ranged from 19 to 69 (*M* = 34.9, *SD* = 10.95), 61% identified as male, 69% identified as White, and no participants indicated that they were Muslim.

Participants first read nine statements that were ostensibly taken from social media, comment sections, and elsewhere on the internet. Three were negative statements about Muslims (e.g., “With all that’s going on, I think it is OK for people to be suspicious of Muslims”), and three were negative statements about politicians (e.g., “All politicians really care about is themselves. They’ll do anything to get more and more power”). I included three control statements that were also negative and went against descriptive norms, but about trivial tastes and not loaded with the same moral weight as prejudiced statements. These statements were about people disliking cookies, the beach, and pizza (e.g., “What’s the big deal about pizza? I’ve never tasted a slice of pizza that tasted good”). In response to each of these statements, participants indicated on a seven-point scale (from 1, *Strongly Disagree* to 7, *Strongly Agree*) how much they thought that, in saying the statement, that person was being “true to themselves,” “honest,” “authentic,” and “genuine.” For each group of statements (Muslims, politicians, cookies, beach, and pizza), these items were averaged together to measure *perceived authenticity*.

Participants then filled out a demographic questionnaire. At the end of this page, I measured prejudice against Muslims and politicians, as well as how much participants disliked cookies, the beach, and pizza.

Anti-Muslim prejudice was measured using seven items adapted from measures of modern prejudice (Biernat & Crandall, 1999) that tapped into beliefs about Muslims (e.g., “Muslims living here should not push themselves where they are not wanted”) as well as feelings toward them (e.g., “How much do you like or dislike Muslims?”).

Anti-politician prejudice was measured using seven items, some adapted from a standard social distance questionnaire (Biernat & Crandall, 1999; “I would like a politician to be a close personal friend”) and others tapping into feelings toward politicians (e.g., “I admire politicians”).

Lastly, participants were asked how much they liked cookies, the beach, and pizza on a seven-point scale (from 1, *Not at all* to 7, *Very much so*). These items were reverse-scored to indicate control dislike statements—ones that are non-normative, but trivial tastes.

**Results**

Correlations that test the primary hypothesis are reported in the diagonal of Table 1; this table presents the correlations between each combination of dislike and perceived authenticity measures. The more prejudice people reported toward Muslims, the more they perceived similarly-prejudiced statements about Muslims to be authentic, *r* = .38, *p* < .001, and the more prejudice people reported towards politicians, the more they thought prejudiced statements about politicians were authentic, *r* = .18, *p* = .048. However, control dislikes did not correlate with thinking the control dislike statements were authentic.

Was this due to a *general* tendency for people who report prejudice to also think negative statements represent people’s authentic selves? EXPLAIN SUCCINCTLY HERE. Confidence intervals for these differences in correlations were calculated using Zou’s (2007) method (Diedenhofen & Musch, 2015). The correlation between prejudice against Muslims and perceived authenticity of anti-Muslim statements was .38, while the correlation between prejudice against Muslims and perceived authenticity of anti-politician statements was .05. The difference between these two was .33 [.16, .50]; the relationship between Muslim prejudice and thinking anti-Muslim statements are authentic were stronger than the relationship between prejudice against Muslims and perceived authenticity of anti-politician statements. This was also the case with anti-politician prejudice and perceived authenticity of anti-Muslim statements, .24 [.06, .41].

Table 1

*Correlations Between Dislike and Perceived Authenticity*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Perceived Authenticity | | | | |
| Dislike | Muslims | Politicians | Pizza | Beach | Cookies |
| Muslims | **.38** | .05 | -.16 | -.05 | -.16 |
| Politicians | -.06 | **.18** | -.02 | .02 | -.05 |
| Pizza | -.13 | -.15 | -.03 | .01 | -.02 |
| Beach | **-.24** | **-.28** | -.10 | -.17 | -.06 |
| Cookies | -.05 | -.12 | .05 | -.09 | .05 |

Bolded *r*s, *p* < .05.

**Study 2**

**Method**

Participants were recruited as they were walking around the University of Kansas campus. Research assistants approached passers by and asked if they would like to fill out a short, one-page survey in exchange for a piece of candy. Sample size was determined by how many people I could recruit by the end of the semester. A total of 221 people participated, but 7 participants were excluded for partial responding.

Participants were randomly assigned to one of two conditions: An *illegal immigrant condition* or a *Kansas State condition*. In the former, participants read two negative statements about illegal immigrants (e.g., “With all that’s going on, I think it is OK for people to be suspicious of illegal immigrants”); in the latter, participants read two negative statements about Kansas State students—a rival school of the University of Kansas (e.g., “Students that go to Kansas State smell weird”). The same measure of perceived authenticity was used in this study as in Study 1, and participants were again told that these statements came from social media websites and comment sections on the internet. Participants then reported how they felt about illegal immigrants and Kansas State students.

Five of the items used in Study 1 for Muslims were employed here, adapted to illegal immigrants (e.g., “Illegal immigrants here teach their children values and skills different from those required to be successful in the United States”); five of the items used in Study 1 for politicians were employed here, adapted to Kansas State students (e.g., “Kansas state students or graduates are similar to me”). In order to keep the survey at a one-page limit, demographic questions were not asked.

**Results**

I predicted that prejudice against illegal immigrants would predict perceived authenticity, but only in the illegal immigrant condition, while anti-Kansas State prejudice would only predict perceived authenticity in the Kansas State condition. These two prejudices were positively correlated, *r* = .28, *p* < .001, so I use the irrelevant prejudice as a control in all analyses.

First, I regressed perceived authenticity on anti-Kansas State prejudice, anti-illegal immigrant prejudice, condition, and an interaction between the latter two predictors. The condition by anti-illegal immigrant prejudice was significant, *b* = -.44, *SE* = 0.18, *t*(209) = -2.46, *p* = .015. Prejudice against illegal immigrants was positively related to authenticity in the illegal immigrant condition, *b* = .36. *SE* = .12, *t*(209) = 2.98, *p* = .003; however, it was not related to authenticity in the Kansas State condition, *b* = -.08, *SE* = .13, *t*(209) = -0.56, *p* = .575.

Next, I regressed authenticity on anti-illegal immigrant prejudice, anti-Kansas State prejudice, condition, and the interaction between the latter two predictors. The condition by anti-Kansas State prejudice interaction was significant, *b* = .45, *SE* = .16, *t*(209) = 2.75, *p* = .007. Prejudice against Kansas State students was positively related to authenticity in the Kansas State condition, *b* = .53, *SE* = .12, *t*(209) = 4.44, *p* < .001; however, it was not related to authenticity in the illegal immigrant condition, *b* = .08, *SE* = .12, *t*(209) = 0.66, *p* = .513.

**Study 3**

**Method**

I recruited 200 people from MTurk to participate in a “survey on perceiving other peoples’ attitudes.” Throughout this paper, participants who completed one of these studies were barred from participating in a subsequent one (Litman, Robinson, & Atterbock, 2017). The analyses for this study involves multilevel modeling, and I was unsure of proper a priori expected population parameters to choose for a power analysis, so sample size was determined from a subjective decision of what seemed appropriate. I recruited 200 participants, and each participant contributed 10 data points; I found a level one *n* = 2000 and a level two *n* = 200 to be reasonable. Participants’ ages ranged from 19 to 70 (*M* = 34.15, *SD* = 11.56), 54% identified as male, and 76% identified as White. Participants answered a questionnaire in the following order.

**Target groups.** Most of the following questions were asked once for each target group. These target groups were: Black people, transgender people, fat people, police officers, lawyers, business people, prostitutes, drug dealers, blind people, and deaf people.

**Perceived descriptive normativity.** Participants were asked to think about Americans in general and indicate “what percentage of Americans, if they were being truly and totally honest with themselves, would admit they feel negatively” toward each of the target groups; participants indicated their responses on a 0 to 100 scale.

**Perceived prescriptive normativity.** Participants were then asked to pivot from thinking about what “Americans *actually* feel” to think about “how Americans think people *should* feel.” They were asked to indicate on the same 0 to 100 scale what percentage of Americans “think it is OK to feel negatively toward these groups?”

**Perceived authenticity.** Ten quotations, ostensibly “taken from social media posts or comments on the internet,” were presented to participants. Each target group had one corresponding negative statement about them (e.g., “Blacks are the people causing the racial tension in America today,” “Business people don’t care about anyone but themselves and making lots of money”); the same four items employed in Studies 1 and 2 were asked for each statement.

**Prejudice.** Participants were asked how they feel about each of the target groups on a scale from 0 (*Very negatively*) to 100 (*Very positively*). These items were reverse-scored such that higher scores indicated more prejudice toward the group.

**Concern about political correctness.** The belief that there are “PC crusaders” in society censoring speech, changing history, and dividing the country with identity politics were measured using six items (Lalonde, Doan, & Patterson, 2000); participants were asked how much they agree with these statements (e.g., “I believe that there are ‘PC’ crusaders who want to censor people’s speech”) on a seven-point scale.

**Demographics.** Lastly, participants indicated their age, gender identity, race, political outlook (1, *Very liberal* to 7, *Very conservative*), and political affiliation (1, *Strongly Republican* to 7, *Strongly Democrat*). Political outlook and reverse-scored political affiliation were averaged together to measure right-wing political identification (*r* = .85).

**Analysis details.** As constructs were measured on very different scales (i.e., seven- versus 101-point), all measures were standardized (across, not within, individuals) before analyses. Measurements were considered at the first level, while individuals were at the second. Thus, perceived descriptive and prescriptive normativity, perceived authenticity, and prejudice were at the measurement level, while political correctness concern and right-wing political identification were at the person level. All regression coefficients, when measured at the first level, were allowed to differ by individual; that is, random slopes were defined for all level-one predictors. Lastly, I used Satterthwaite’s approximation for degrees of freedom for all *t*-tests of regression coefficients (Kuznetsova, Brockhoff, & Christensen, 2017).

**Results**

Prejudice again correlated positively with perceived authenticity of negative statements, *b* = .24, *SE* = .03, *t*(181.53) = 8.31, *p* < .001. Additionally, the *more* that participants thought others had the prejudice (i.e., descriptive normativity), the *more* authentic they perceived the speaker to be, *b* = .20, *SE* = .02, *t*(186.24) = 8.60, *p* < .001.

**Moderator analyses.** I tested if the relationship between prejudice and perceived authenticity was moderated by (a) perceived prescriptive normativity, (b) concern about political correctness, or (c) right-wing political identification. For each of these moderators, I regressed perceived authenticity on prejudice, the moderator, and the interaction between the two. Each of the three models included prejudice as a random slope; when testing the interaction by prescriptive normativity, coefficients for both prescriptive normativity and the interaction were included as random slopes, as well.

The prejudice by prescriptive normativity interaction on authenticity was significant, *b* = -.06, *SE* = .02, *t*(156.39) = -2.82, *p =* .005. Probing this interaction (Preacher, Curran, & Bauer, 2006) showed that, when participants reported *low* perceived prescriptively normativity (i.e., a standard deviation below the mean), the relationship between prejudice and authenticity was *b* = 28, *SE* = .05, *z* = 6.13, *p* < .001. When the group was high perceived prescriptive normativity (i.e., a standard deviation above the mean), the relationship was about half as strong, *b* = .15, *SE* = .03, *z* = 4.96, *p* < .001.

The prejudice by political correctness concern on authenticity was *not* significant, *b* = .005, *SE* = .03, *t*(172.5) = 0.17, *p* = 0.864.

**Study 4**

**Method**

I recruited 200 participants from MTurk to participate in a “study on person perception.” As this study aimed to eliminate the correlation between prejudice and perceived authenticity with an experimental manipulation, sample size was determined by simulating data where the correlation between two variables was *r* = .40 for half of the participants and *r* = .00 for the other half, then choosing the sample size that led to 80% power. This power analysis was used to inform sample sizes in all subsequent studies. Two participants failed to complete the writing task (described below); they were excluded from all analyses, leaving a final sample size of 198. Participants’ ages ranged from 19 to 77 (*M* = 36.39, *SD* = 11.80), 52% identified as male, and 76% identified as White.

Participants were told that the study was aimed at the question, “Why do we think that others are the way they are?” They were told that they would answer a few questions and then comment on previous participants’ responses to those same questions. Participants were then randomly assigned to an *expression condition* or a *suppression condition*.

In both conditions, participants were told that they were in the version of the survey about “fat people,” and they were asked to give a number of reasons why people might be fat. They were given ten blank lines below the instructions to do so. In the *expression condition*, they were told that it was “important that you feel free to write whatever reasons” that they think of, whether they agree with them, whether they think the reasons are nice or mean. In the *suppression condition*, they were told that it was important that their reasons “are not focused on blaming fat people for their bodies,” because “quite a lot of research shows that blaming people for their weight is a sign of prejudice.” In this latter condition, an additional question asked them to look over their answers again, making sure that nothing they said blamed fat people for their weight. After double-checking, they were instructed to select a button that read, “Yes, I followed the directions.”

On the next page, participants were told: “Some people already answered the *exact same* question that you just answered. Here is one of the reasons that they gave for people having obesity...” This was repeated four times, each with a new statement. Two blamed fat people for their weight (i.e., “they have no willpower,” and “they’re too lazy to exercise”)—I refer to these as the *negative statements*. The other two—the *neural statements*—did not (i.e., “their genes make them overweight,” and “environmental things like poverty or bad parenting”). Participants were asked five questions about each statement.

**Perceived authenticity.** The same four questions used in Studies 1 – 3 were used to measure the perceived authenticity of the speaker. The eight items for the negative statements were averaged together, while the eight items for the neutral statements were averaged together to measure perceived authenticity.

**Manipulation check.** Participants were asked how much they agreed with the statement, “This answer follows the rules of the task,” on a seven-point scale (1, *Strongly disagree* to 7, *Strongly agree*). The items for the positive and neutral statements were again averaged together separately.

Participants then answered a demographics questionnaire, followed by Crandall’s (1994) anti-fat attitudes questionnaire. In line with the theory behind the scale, I used the “Dislike” subscale as my measure of anti-fat prejudice. An example item reads: “I really don’t like fat people much.” Participants indicated how much they agreed with these statements on a seven-point scale (1, *Strongly disagree* to 7, *Strongly agree*).

**Results**

The two negative statements were seen as following the rules less in the suppression condition (*M* = 1.89, *SD* = 1.27) than in the expression condition (*M* = 6.1, *SD* = 0.85), *t*(158.31) = 27.12, *p* < .001, *d* = 3.86 [3.39, 4.43]. There was no difference between the two conditions for the neutral statements, *t*(170.52) = 1.08, *p* = .281, *d* = 0.15 [-0.13, 0.43]. Note that Welch’s *t*-test was employed due to unequal variances across conditions (Delacre, Lakens, & Leys, 2017).

I tested my hypothesis by regressing perceived authenticity of the negative statements on anti-fat prejudice, condition, and the interaction between the two. The interaction was significant, *b* = .36, *SE* = .12, *t*(194) = 3.16, *p* = .002. Probing the interaction with simple slopes analyses (Preacher, Curran, & Bauer, 2006) showed that prejudice was positively correlated with perceived authenticity in the suppression condition, *b* = .34, *SE* = .08, *t*(194) = 4.41, *p* < .001. There was no relationship between the two in the expression condition, *b* = -.02, *SE* = .08, *t*(194) = -0.23, *p* = .815.

This interaction held even after including the other two anti-fat attitudes subscales—fear and willpower—as additional predictors, *b* = .21, *SE* = .11, *t*(192) = 1.99, *p* = .048.Regressing the perceived authenticity of the neutral statements on anti-fat prejudice, condition, and the interaction between the two yielded a nonsignificant interaction, *b* = .10, *SE* = .10, *t*(194) = 1.05, *p* = .293.

**Study 5**

**Method**

I recruited 200 people from MTurk to participate in a “study on person perception.” A total of 202 people participated, but one was dropped from analyses for partial responding. Participants’ ages ranged from 18 to 70 (*M* = 36.71, *SD* = 11.70), 42% identified as male, and 77% identified as White.

The procedure was identical to Study 4, and demographics and prejudice were measured the same as in Study 4. Only the measures after reading both negative and both neutral statements changed. After reading each statement, participants indicated on a seven-point scale how “authentic and genuine,” “true and honest to themselves,” “politically correct,” and “overly careful and too polite” they believed the person was being. The former two items were designed to measure perceived authenticity, with the latter two measuring perceived political correctness.

Considering items in response to negative and neutral statements separately, an exploratory principal axis factor analysis with an oblimin rotation supported this two-factor solution and with simple structure (using Kaiser-Guttman criteria, a parallel analysis, and examining scree plots). The four authenticity items for the negative statements were averaged together, while the four items for the neutral statements were averaged together to measure perceived authenticity. The same was done for perceived political correctness.

**Results**

The primary hypothesis was again tested by regressing perceived authenticity of the negative statements on anti-fat prejudice, condition, and the interaction between the two. The interaction was not significant in this study, *b* = .14, *SE* = .12, *t*(197) = 1.14, *p* = .255. However, the simple slopes followed the same pattern as in Study 4: Prejudice predicted perceived authenticity in the suppression condition, *b* = .27, *SE* = .09, *t*(197) = 3.11, *p* = .002, but not in the expression condition, *b* = .14, *SE* = .08, *t*(197) = 1.73, *p* = .085. Regressing perceived authenticity of the neutral statements on prejudice, condition, and their interaction yielded a nonsignificant interaction as in Study 4, *b* = .03, *SE* = .12, *t*(197) = 0.24, *p* = .808.

I also predicted that perceived political correctness of negative statements would be negatively related to perceptions of their authenticity; this was not the case, *r* = -.02, *p* = .818. Regressing perceived political correctness of the two negative statements on anti-fat prejudice, condition, and their interaction did not yield a significant interaction, *b* = -.07, *SE* = .13, *t*(197), *p* = .609. There was a significant zero-order correlation between prejudice and perceived political correctness, however, *r* = .30, *p* < .001.

**Study 6**

**Method**

I recruited 230 people from MTurk to participate in a task involving “classifying statements.” A total of 229 people participated; their ages ranged from 19 to 63 (*M* = 34.13, *SD* = 9.74), 55% identified as male, and 77% identified as White.

Participants were told: “We are creating a statistical model that will be able to predict how people will perceive others. These models must first learn from humans, however; this is where we need your help! In this task today, you will be reading statements made by others and classifying them...” They were told that the specific concept of focus was “*authenticity*,” and that they would be (a) classifying statements *about* authenticity as “good” or “bad” and (b) reading statements taken from the internet and rating how authentic those people were being in expressing those statements.

After indicating that they had read the instructions carefully, participants were randomly assigned to either an *authenticity is good* or an *authenticity is bad* condition. In the former, participants were given eight positive (e.g., “Being authentic can be soothing because it reduces stress in your life. If you are acting in a natural way without trying to change your behavior that is one less thing to worry about”) and two negative (e.g., “Sometimes being authentic means ignoring the expectations and desires of people around you. If you always act the way you truly feel deep down, you can hurt people's feelings and make them uncomfortable”) statements to rate as “good,” “bad,” or “unsure.” In the latter, participants were given two positive and eight negative statements.

On the same page, participants were asked, “How much of a good thing do you think being ‘authentic’ is?” and “How good of a thing do other people think being ‘authentic’ is?” on a scale from 1 (*Extremely bad*) to 7 (*Extremely good*). These items served as a manipulation check.

Participants were given a list of ten statements and asked to rate how authentic they thought people were being, on a scale from 0 (*Not at all*) to 100 (*Very much so*). Two of the statements had to do with illegal immigrants: “With all that’s going on, I think it is OK for people to be suspicious of illegal immigrants,” and “Illegal immigrants threaten what it means to be American. It is STUPID to let them come flooding in.” These two items were averaged together to represent perceived authenticity of prejudiced statements. The other eight statements were statements one might see elsewhere on the internet (e.g., “I’m pulling for the Minnesota Twins this year, they’re a fun team”).

Participants then filled out a demographics questionnaire and then indicated on a seven-point scale how much they agreed with a number of statements that were related to what they rated as authentic. Three were items illegal immigrants used in Study 2 (e.g., “Undocumented aliens here teach their children values and skills different from those required to be successful in the United States”), which were averaged together to form a measure of prejudice against illegal immigrants.

**Results**

Participants rated the goodness of authenticity higher in the *authenticity is good* condition (*M* = 6.30, *SD* = 0.82) than in the *authenticity is bad* condition (*M* = 5.28, *SD* = 1.23), *t*(198.53) = 7.38, *p* < .001, *d* = 0.98 [0.70, 1.25]. They also indicated that they believed *others* thought authenticity was better in the *authenticity is good* (*M* = 6.14, *SD* = 0.79) condition than in the *authenticity is bad* (*M* = 3.29, *SD* = 1.46) condition, *t*(175.14) = 18.42, *p* < .001, *d* = 2.43 [2.09, 2.78]. Welch’s *t*-test was again used due to unequal variances across conditions (Delacre et al., 2017).

The primary hypothesis was tested by regressing perceived authenticity on prejudice, condition, and the interaction between the two. The interaction was not significant, *b* = 1.74, *SE* = 2.49, *t*(225) = 0.70, *p* = .486. However, I replicated the same correlation between perceived authenticity and prejudice, *r* = .21, *p* = .002.

**Study 7**

I recruited 220 people from MTurk to participate in a “survey on perceiving others.” Participants’ ages ranged from 18 to 71 (*M* = 36.24, *SD* = 11.55), 47% identified as male, and 75% identified as White.

Participants were asked to “read about someone and answer questions about something they said” and randomly assigned to either an *accuracy* motivationor *expression* motivation condition. In the accuracy condition, I asked participants to “be as accurate as possible” and to “try to guess the correct answer” when reporting their perceptions of the target person. In the expression condition, I asked participants to “answer based on expressing your own opinion” and to “respond in a way that expresses what you personally think.” To help participants adhere to these goals, participants were told that they would be getting a bonus payment upon completion and that, “We only ask in return that you [be accurate in your ratings/respond based on your own opinion].”

Participants then read about a person named Colin, as well as some information about him: where he lives, what he does for work, food he likes, and some basic personality characteristics (also see White & Crandall, 2017; White & Molina, 2016). This was meant to be general, somewhat bland information for the purpose of allowing participants to feel like they have sufficient background information to reason with and make judgments about the target. At the end of the description, participants were told that Colin recently said the following statement: “Black people are so touchy about race that it is difficult to get along with them. They can be combative and assume the worst from White people. This makes me feel uncomfortable sometimes, which is why I don’t really like to hang around them much.”

I reminded participants about their goal before measuring perceived authenticity using the same items as in Studies 1 – 4. Participants answered a demographics questionnaire. Prejudice was then measured using an eight-item, seven-point symbolic racism scale (Henry & Sears, 2002; “Irish, Italian, Jewish, and many other minorities overcame prejudice and worked their way up. African-Americans should do the same, without any special favors”).

**Results**

I tested the primary hypothesis by regressing perceived authenticity on prejudice, condition, and their interaction. The interaction term was not significant, *b* = .09, *SE* = .13, *t*(216) = 0.69, *p* = .491. However, the simple slopes showed the predicted pattern: prejudice positively predicted authenticity in the expression condition, *b* = .25, *SE* = .10, *t*(216) = 2.57, *p* = .011, but not in the accuracy condition, *b* = .15, *SE* = .09, *t*(216) = 1.65, *p* = .101.