**Perceived Authenticity as a Justification for Prejudice**

Donald Trump has habitually expressed prejudice throughout his campaign and into his presidency. He has openly insulted Mexicans, immigrants, Muslims, Haitians, Africans, Black Americans, Puerto Ricans, Korean Americans, women, and people who are overweight (Lee & Quealy, 2018; Leonhardt & Philbrick, 2018). Prejudice has plagued politics in the United States since the country’s inception (e.g., Anderson, 2017; Kendi, 2017), but Trump’s expressions are the most *explicit* seen from a presidential candidate in decades (Haney-López, 2013; Omi & Winant, 1994).

Americans’ reactions to this rhetoric ranged from abhorrence to admiration. One such reaction was to label Trump as “authentic,” “genuine,” “telling it like it is,” and “speaking his mind.” People from across the political spectrum—conservative to progressive—called Donald Trump “*the* authentic” candidate (Estepa, 2017; Johnson, 2017; Merelli, 2017); polling suggests many Republican primary voters thought the same (Sargent, 2015). A number of factors likely led to people seeing Trump as “authentic,” but his flouting of norms against explicitly expressing prejudice has been one of the oft-mentioned reasons for so doing (Basavaraju, 2016; Bebout & Garcia, 2016; Gallagher, 2017; Johnson, 2016).

I examine this phenomenon across eight studies, generalizing beyond Donald Trump: Why do people perceive another’s expression of prejudice to be authentic? I propose that people will perceive prejudice as authentic when they hold that same prejudice. In the following sections, I discuss the nebulous concept of authenticity—both how theorists conceptualize it as well as how people perceive it—before examining how this concept fits within modern theories of prejudice expression. I then propose a number of hypotheses that may explain *why* self-reported prejudice might predict greater perceived authenticity of that same prejudice.

**A Brief History of the Psychological Study of Authenticity**

Philosophers have been interested in the concept of being authentic or true to oneself for thousands of years, but the psychological study of authenticity was first embraced by humanistic psychologists (Kernis & Goldman, 2004). In this tradition, authenticity was vaguely defined as living in accordance with one’s true, core, inner self. These psychologists saw the inner core of people as being fundamentally good and living in accordance with this core to be instinctual, so they believed the job of the therapist was to *facilitate* clients living in accordance with their inner selves. Many ideas in the humanist tradition were treated synonymously with “authenticity,” even if they were not always labeled as such.

*Congruence* was the fundamental concept in the way Carl Rogers saw psychotherapy. The idea was to align how a client thinks about themselves with their ideal version of themselves. Rogers (1961) quoted Kierkegaard to exemplify the goal he set for his clients: “to be that self which one really is.” Although he did not explicitly call this authenticity, it matches with contemporary definitions of the term, and he places centrality on closely-related concepts like being genuine, true, real, and not “putting up facades” to others. Abraham Maslow’s work focused on *self-actualizing*, which he defined as an “acceptance and expression of the inner core or self” (Maslow, 1968, p. 197). Again, this is similar to modern notions of authenticity. He touched on the concept explicitly by arguing an authentic person is one who resists influences to deviate from their inner core (Maslow, 1968).

Sidney Jourard defined authenticity as something fundamentally social. He argued that we learn at a young age, through punishments and rules, to suppress our true feelings. If done chronically and inappropriately, this leads to negative psychological functioning. He saw *self-disclosure* of the real self to others—being *transparent*—as vital for well-being. He defines the authentic person as one who is “being oneself, honestly, in one’s relations with his [sic] fellows” (Jourard, 1964, p. 153). Much of his work reiterates that one of the main goals of therapy is to help the client live an authentic life—for the client to let people see them how they see themselves and resist the suppression of their true self by external pressures.

These humanistic ideas about authenticity were not subjected to many empirical tests, however, partly due to the elusiveness of determining what precisely is one’s true self—“it is unclear what the self is being true to, when it is being true to itself” (Sheldon, 2009, p. 75). Positive psychologists (Sheldon & King, 2001) study authenticity by doing away with the idea that there is an *objective* inner self, given that the self is socially constructed, changing with one’s goals, contexts, and social roles. Sheldon, Ryan, Rawsthorne, & Ilardi (1997) defined authentic behavior as that which is “*phenomenally* experienced as being authentic by the self… or internally caused,” and that to act authentically is to act “with a full *sense* of choice and expression” (p. 1381, emphasis mine). Authenticity is thus a subjectiveexperience. This definition is remarkably similar to the psychological need for autonomyin the context of self-determination theory, where autonomy “connotes an inner endorsement of one’s actions” and “people experience themselves as initiators of their own behavior” (Deci & Ryan, 1987, p. 1025). Research conducted under the umbrella of self-determination theory has found consistent positive consequences for *feeling* authentic, such as increased happiness and well-being (e.g., Kernis & Goldman, 2004; Ryan & Deci, 2004; Sheldon & Elliot, 1999).

**Perceived Authenticity**

There are many academic definitions for authenticity (Sheldon, 2009), and many researchers note that it is often defined vaguely or not at all in the literature (e.g., Beverland, 2005; Kadirov, 2015; Molleda & Jain, 2013). But the present research question concerns how the lay person perceives authenticity. How do people define and judge authenticity? How do perceptions of authenticity affect attitudes and behavior? Evidence answering each of these questions is demonstrated across various fields of study, such as psychology, sociology, marketing, and communications.

There is no one definition for authenticity. What is perceived as authentic depends on the person and the context; perceptions of authenticity are a “socially constructed interpretation… of what is observed rather than properties inherent in the object” (Beverland & Farrelly, 2010, p. 839; see also Grayson & Martinec, 2004). Qualitative research best shows the breadth of ways people understand authenticity.

Beverland and Farrelly (2010) asked participants how they defined authenticity, presenting them with about 100 images depicting various brands, cultural icons, tourist sites, etc., to facilitate discussion about what they perceived to be authentic. Among many other reasons, interviewees attributed authenticity to products because the product did what it was advertised to do (shampoo) or gave people greater ability to develop knowledge (Apple computers), while they ascribed inauthenticity to products for being expensive and exploitatively produced (Nike) or being unhealthy (cigarettes).

Items in museums can literally be authentic in that they are *actual* historical objects from an important time or relating to an important figure (Grayson & Martinec, 2004; Hede, Garma, Josiassen, & Thyne, 2014). Interestingly, however, people are very willing to assign authenticity to objects that they *know* are contrived attempts to represent a fiction. Grayson and Martinec (2004) interviewed people visiting Shakespeare’s birthplace and the Sherlock Holmes museum, finding that people used different definitions of authenticity in their judgments in labeling each of these attractions as authentic. The authors define *indexical authenticity* to refer to an object that is the original—not a copy or replica—while *iconic authenticity* refers to an object that accurately represents the original, without actually being the original. For example, an indexically authentic basketball jersey is one a player wore while playing in a game, while an iconically authentic jersey is one made *exactly* like the player’s, but was never actually game worn.

This concept of iconic authenticity was exemplified by patrons of the museum finding Sherlock Holmes’s—a fictional character—“possessions” to be authentic because: “You have a feeling as if [Holmes] really touched everything,” the museum contained “things that Sherlock Holmes might have used to solve the crimes,” and that one could see “[Holmes’s] living quarters, and where he sat, did his writings and thinking” (pp. 300-301). Patrons also noted the authenticity of the Holmes museum because everything inside looked as if it belonged at that historical time; in this case, people’s perceptions of authenticity are tangled up with feelings of nostalgia for a (fictionalized) past.

Kovacs, Carroll, and Lehman (2014) asked participants to choose or nominate words that would describe a restaurant as authentic. In addition to the synonyms commonly found across authenticity research (e.g., genuine, real), participants also associated authentic restaurants with words like skilled, traditional, historical, expert, professional, and iconic. Some of these words contradict what people perceive to be authentic in political candidates, as amateurism and non-professionalism can lend an aura of authenticity to a candidate (Enli, 2017; Manning, Penfold-Mounce, Loader, Vromen, & Xenos, 2017).

Researchers across disciplines have also studied the correlates, causes, and consequences of perceived authenticity—a literature that is not tied together by an overarching theoretical framework. Psychologists have found that perceiving romantic partners as authentic relates to greater relationship quality (e.g., increased satisfaction, less avoidance; Wickham, 2013), seeing one’s coworker as authentic relates to liking and trusting that coworker (Liu & Perrew, 2006), and people high in power (i.e., a CEO) are seen as less authentic in expressing how they feel than those in less power (Kim et al., 2017).

Marketing research also demonstrates positive consequences of perceived authenticity. The more people perceive a variety of products to be authentic, the more they are willing to pay for them (Kadirov, 2015). Kovacs, Carroll, & Lehman (2014) analyzed nearly 19,000 Yelp reviews for restaurants, finding that the more that people mentioned authenticity-related words, the more positively they reviewed the establishment. Family-owned restaurants were perceived as authentic, while chain restaurants were seen as inauthentic. Restaurants could also be categorized into keywords, such as Italian, Chinese, burgers, sandwiches, pizza, etc., and the more categories restaurants belonged to, the less authenticity-related words reviewers used. Kovacs and colleagues replicated these findings in a vignette experiment.

Believing that one’s Airbnb stay was authentic (i.e., representative of local way of life, conducive to interacting with the local community) was positively related to consumers agreeing they paid a reasonable price and that the stay was a good value (Lian, Choi, & Joppe, 2016). Public figures are perceived as less authentic when they hire someone to write their social media posts for them (Cohen & Tyler, 2016). In-depth interviews with business owners and political consultants show that authenticity is a concept thought to be important in brand management (Beverland, 2005; Sorazio, 2017).

Two studies on perceived authenticity are more directly related to the current research question. Hahl, Kim & Sivan (2017) examined how people could find a “lying demagogue” to be such an authentic political candidate. In a minimal-group experiment, Hahl and colleagues found that participants who read about a candidate telling obvious lies and making misogynist remarks perceived this candidate to be *more* authentic when the participant felt like they were from an aggrieved social group and that the political system was not representing their interests. The authors argued that this was because lying and misogyny flout the established norms of a political system that participants found illegitimate.

Pillow, Crabtree, Galvin, and Hale (2017) also investigated perceived authenticity of politicians, arguing that people are motivated to see those they support as authentic. In a survey about five candidates from the 2016 United States presidential election, they found that the positive relationship between perceiving the candidate to speak candidly and perceiving the candidate to be authentic was stronger when people liked the candidate (compared to when they did not). Pillow and colleagues argued that this supports a motivated reasoning (Kunda, 1990) account; they interpreted the interaction to be due to people liking the candidate weighting perceived candor higher when judging the authenticity of a candidate then those who dislike the candidate. The authors also demonstrated that perceiving the candidate as unfiltered and not “politically correct” predicted people thinking the candidate was rejecting external influences on their behavior, which in turn predicted perceiving the candidate as authentic. They argued that this supports a correspondent inference (Jones & Davis, 1965) account, given that people assume others are acting on their own volition when when acting against norms.

**Defining Perceived Authenticity**

I define *perceived authenticity* as an observer’s judgment of how much a target individual is acting in accordance with what the observer perceives to be the target’s actual beliefs, attitudes, goals, and desires (also see Liu & Perrew, 2006). This situates the construct wholly in the observer. For example, Daisy believes Adam supports Democrats, while John believes Adam supports Republicans; if Daisy and John see Adam at a rally for a Democratic candidate, Daisy will perceive Adam’s actions to be authentic, while John will see them as inauthentic. Whether or not Adam actually supports Republicans or Democrats is irrelevant.

**Justifying Prejudice: The Justification-Suppression Model**

Prejudice is a negative evaluation of a social group or of an individual primarily based on their presumed group membership (Crandall & Eshleman, 2003). People feel and want to express prejudice, yet it is socially unacceptable to do so. The justification-suppression model (JSM; Crandall & Eshleman, 2003) models this tension, synthesizing the commonalities between various modern theories of prejudice into a process with four components: genuine prejudice, suppression, justification, and expression.

*Genuine prejudice* is the unmanaged, unalloyed underlying negative affect one feels toward a social group or member of a social group; it is the motivational state (Brehm, 1999) that drives justification and expression. Unfortunately, genuine prejudice is considered to be unmeasurable, as any self-report will be contaminated by biases (e.g., social desirability) and one’s inability to report feelings of which they are not aware. Implicit and physiological measures are promising techniques for tapping into genuine prejudice, but there are still unresolved questions about what precisely these measures capture, given that they tend to correlate weakly with one another (Fazio & Olson, 2003) and do not demonstrate greater predictive validity than do self-reports (Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013).

*Suppression* is an externally or internally motivated (Plant & Devine, 1998) attempt to inhibit the expression of one’s prejudice. The primary external motivators are social norms and institutional rules (e.g., workplace code of conduct) that proscribe the expression of prejudice. Internal motivators can be one’s values and beliefs that conflict with expressing prejudice (e.g., egalitarianism) or one internalizing external motivators (Deci & Ryan, 2000). Suppressive forces do *not* necessarily eliminate the underlying genuine prejudice—they merely keep genuine prejudice from being expressed. Crandall and Eshleman (2003) review a wide variety of research demonstrating that suppressing prejudice is taxing on an individual; it requires cognitive effort. People are motivated to relax this effort and free the expression of genuine prejudice. To do so, people seek out justifications for prejudice.

*Justifications* are any social or psychological process that allows one to feel as if they can express prejudice without suffering negative external (e.g., reprimand from others; workplace punishment) or internal (e.g., guilt, shame) consequences. These justifications release the *expression* of prejudice. For example, when people have already demonstrated that they can be non-prejudiced, they can feel justified to act in discriminatory ways (Monin & Miller, 2001); people express prejudice when they feel like they have demonstrated that they have standing to do so (Miller & Effron, 2010).

**Authenticity as Justification for Prejudice**

I propose that people who are prejudiced are more likely to perceive an expression of prejudice as “authentic.” One can see this as a justification for prejudice by considering *vicarious justification* and *justification by rearticulation.*

**Vicarious justification.** Anecdotal evidence abounds that people will justify another’s expression of prejudice. This is clear from studying the campaign and presidency of Donald Trump (e.g., Bump, 2018; Burton, 2018; Horsey, 2016; Scott, 2018). White and Crandall (2017) demonstrated this empirically. Across a series of experiments, participants were told that a man was fired for expressing (a) anti-Black prejudice or (b) control statements, such as negative statements about police or coworkers. Participants were then asked how much they agreed that firing the employee violated his freedom of speech (e.g., “[The employee’s] bosses disrespected his right to free speech”). Various measures of prejudice predicted relevance of freedom of speech in the anti-Black prejudice conditions (meta-analytic *r* = .43) but not in the control conditions (meta-analytic *r* = .09). Participants were *not* principled in marshaling freedom of speech, showing people defend others’ prejudiced expressions as a function of their own. White and Crandall found that this relationship was partially due to the termination of a prejudiced employee threatening the expressive autonomy (e.g., “I feel free to express my ideas and opinions”) of prejudiced participants. People engage in vicarious justification: They personally feel the suppression placed on similarly-prejudiced others and strategically deploy values to protect these others. Labelling another’s expression of prejudice as authentic may be a way to vicariously justify prejudice.

**Justification by rearticulation.** Omi and Winant (1994) argued that racism did not necessarily decrease in the United States after the gains of the Civil Rights Movement; their focus is less on if racism is “getting better” and more on how racist ideology takes on new forms. They argue that racism is *rearticulated*: “The new right generally does not display *explicit* racism. It has gained political currency by rearticulating racial ideology” (p. 127). Omi and Winant discuss how conservative racial ideology in the United States rearticulated from overt racism to ostensibly race-neutral concepts. Lee Atwater, a then aide to Republican president Ronald Reagan, explicitly described his conscious effort to rearticulate racial arguments:

You start out in 1954 by saying, “Nigger, nigger, nigger.” By 1968 you can’t say “nigger”—that hurts you, backfires. So you say stuff like, uh, forced busing, states’ rights, and all that stuff, and you’re getting so abstract. Now, you’re talking about cutting taxes, and all these things you’re talking about are totally economic things and a byproduct of them is, blacks get hurt worse than whites…. “We want to cut this,” is much more abstract that even the busing thing, uh, and a hell of a lot more abstract than “Nigger, nigger” (Perlstein, 2012).

Although Omi and Winant focus on Reagan-era conservatism, liberal politicians have also been guilty of rearticulating racism (Haney-López, 2013).

Prejudice and perceived authenticity may be positively related because claiming prejudice as being “authentic” is a rearticulated way of expressing prejudice. It is a justification in a sense that it allows someone to express that they agree with the prejudiced statement in an ostensibly race-neutral way—that is, without ever explicitly mentioning the target group.

**Contributing Mechanisms**

There are a number of psychological phenomena, beyond simple agreement, that could explain the process by which people come to see the expression of prejudice as authentic. I propose four possible accounts: prescriptive norms, social projection, balance, and motivated reasoning.

**Prescriptive norms.** Crandall and colleagues see social norms as the primary suppressive forces acting on the expression of prejudice (Crandall, Eshleman, & O’Brien, 2002; Crandall, Ferguson, & Bahns, 2013; Crandall & Stangor, 2005). *Prescriptive* norms (i.e., norms describing what people *should* do) may influence the relationship between prejudice and perceived authenticity of similarly-prejudiced statements through the aforementioned process of vicarious justification.

People do not need to justify their prejudices when suppression is absent; one needs no justification in saying they hate rapists. Prescriptive norms suppressing the expression of prejudice should motivate justificationsfor prejudice, because these prescriptive norms are threatening to those who hold the unacceptable prejudices. If labelling the expression of prejudice as authentic is a justification, then the positive relationship between prejudice and perceived authenticity should only be present when the prejudice is portrayed as prescriptively non-normative (i.e., suppression is present). This prescriptive norms account argues that perceived authenticity is a mixture of: “I agree with that,” and, “I feel like I cannot express it myself.”

**Social projection.** People believe others are similar to them—an automatic cognitive heuristic termed *social projection* (Kreuger, 2007). The more prejudiced somebody is, the more common they think it is in society; for example, Watt and Larkin (2010) found that participants high in prejudice estimated 71% of people in their country would also be prejudiced; lowly prejudiced participants estimated this to be 51% (see also Pedersen, Griffiths, & Watt, 2008). If prejudiced people think that prejudice is common in society, then they should be more likely to think that any given member of this society is prejudiced. Expressing prejudice is thus perceived as more authentic. In contrast to the JSM, this social projection explanation for the positive relationship between prejudice and perceived authenticity is an unmotivated, cognitive one—it is a simple perceptual bias due to perceiving others are similar to oneself. The social projection account implies two hypotheses: First, participants’ own prejudice should positively predict perceived *descriptive* normativity in society (i.e., how many people *actually* feel prejudice); second, perceived descriptive normativity should *increase* perceived authenticity of prejudiced statements.

**Balance.** Perhaps one of the simplest yet most elegant theories in social psychology is *balance theory* (Heider, 1958, Chapter 7). A primary contention in balance theory is that people are motivated to create a balanced state—a “harmonious state, one in which the entities comprising the situation and the feelings about them fit together without stress” (Heider, 1958, p, 180). A classic example is the “*p-o-x triad*,” in which relations between these three elements *p*, *o*, and *x* can be positive or negative. A balanced state is achieved when the signs of each of the three relations multiply out positively. For example, Esch (1950) presented participants with a situation: A man named Bob (*p*) finds Jim (*o*) to be unintelligent; however, one day Bob reads poetry (*x*) he really likes, tracks down who wrote the poems, and finds out that it was Jim who wrote them. This situation is imbalanced: Bob has a negative opinion of Jim, a positive opinion of Jim’s poems, and Jim has a positive relationship with his poems, given that he wrote them. This multiples out negatively (i.e., -++ = -), thus creating imbalance. Esch asked participants to describe what would happen after Bob realized Jim wrote the poems. About 82% of the participants resolved the imbalance by changing the sign of one of the unit relations. For example, 46% of participants wrote some form of “Bob changes his mind about Jim.” (Readers might be anecdotally familiar with imbalanced states, such as the discomfort of having positive feelings toward two friends who dislike one another, or only liking one member of a married couple.) People searching for affective consistency has been the impetus for a large amount of research in social psychology, and many psychological theories are consistent with balance theory’s predictions (Crandall, Silvia, N’Gabala, Tsang, & Dawson, 2007; Zajonc, 1960).

Consider the present scenario: A participant’s (*p*) relationship with a prejudiced statement (*o*) is positive if they share the same prejudice, and the concept of authenticity (*x*) has a positive, moral connotation (it is virtually axiomatic across all fields that authenticity is generally a *good* thing). To achieve balance, this prejudiced person concludes the prejudiced statement was authentic, creating a positive unit relationship between the two and ensuring affective balance. Balance theory implies the following hypothesis: Presenting authenticity as a *negative* thing should *decrease* the positive relationship between self-reported prejudice and perceived authenticity.

**Motivated reasoning.** Pillow et al. (2017) found that the correlation between perceptions of a political candidate’s candor and perceived authenticity was *weaker* for those that did not like the candidate than for those who did. They argue this is evidence for motivated reasoning, because the information people used to make their judgments of authenticity depended on their pre-existing attitudes. I test this account in a more direct way.

Kunda (1990) argued that motivations can bias cognitive reasoning strategies—people engage in *motivated reasoning.* Kunda saw the motivations stemming from two broad categories: accuracy goals and directional goals. One has a goal for *accuracy* when they are motivated to be correct; people examine information they believe is relevant to the reasoning task at hand, spend more effort processing this information, consider alternatives, and engage with the evidence thoroughly. “Reasoning tasks” are a broad category of occasions that call for one to reason: evaluating a scientific claim, making a decision on who to vote for, impression formation, etc.

One has a *directional* goal when they are motivated to arrive at a particular conclusion. People do not feel comfortable blatantly lying to themselves, so they will still search for evidence—but they will do so in a biased fashion. People do not look for the best information dispassionately and instead focus on evidence, information, reasoning rules, memories, etc., that support their desired conclusion.

Motivated reasoning can account for a number of processes in the domain of stereotyping and prejudice (Kunda, 1990; Kunda & Sinclair, 1999; Sherman, Stroessner, Conrey, & Azam, 2005; Stangor & Ford, 1992). In this research, the most common motivator presumed to shape directional goals is the need for positive self-regard. Stereotyping and prejudice allows people to form favorable comparisons and feel better about themselves. In accordance with the JSM, however, the goal in the present situation is to express prejudice; Crandall and Eshelman (2003) argue that motivated reasoning is a cognitive process that supports the finding of justifications. This implies the following hypothesis: The relationship between prejudice and perceived authenticity should be *weaker* when participants are motivated to be accurate than when they are motivated to express themselves.

**The Present Studies**

Eight studies test six hypotheses. First, self-reported prejudice positively predicts perceived authenticity of prejudiced statements (Studies 1 – 3, 5 – 8). Second, self-reported prejudice should *uniquely* predict prejudiced statements against the *same* target group; conversely, there is no relationship between prejudice and the authenticity of statements in general (Studies 1 and 2). Third, prejudice positively predicts perceived descriptive normativity, which in turn causes people to perceive prejudiced statements as authentic (Studies 3 and 4). Fourth, the positive relationship between prejudice and perceived authenticity is only present when the prejudice is prescriptively non-normative (Studies 3, 5, and 6). Fifth, prejudice should predict perceived authenticity more weakly when authenticity is portrayed negativelythan when it is portrayed positively (Study 7). Sixth, participants motivated to be accurate should display a weaker relationship between prejudice and perceived authenticity than those who are motivated to express themselves (Study 8).

**Study 1**

I tested the basic hypotheses correlationally in this first study. I measured prejudice toward two target groups (Muslims and politicians) as well as perceived authenticity of prejudiced statements against these groups. I predicted that self-reported prejudice would positively correlate with perceived authenticity, but only *within* the same target group (e.g., prejudice against Muslims should not predict perceived authenticity of anti-politician remarks). I measured how much people disliked control targets (the beach, cookies, and pizza) and perceived authenticity of negative statements about these targets, as well. These negative statements were “controls” in that they were negative, descriptively non-normative statements; importantly, however, they lack the same moral implications that prejudices possess. I predicted that there would be no relationship between dislike and perceived authenticity for these control targets.

**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? No—none of the off-diagonal correlations are significant and positive, which would be the case if this general tendency was present. All the off-diagonal correlations were smaller than *r* = +.05, as well. As a formal hypothesis test, I first calculated the difference between the anti-Muslim prejudice correlation with perceived authenticity of anti-Muslim (*r* = .38) and anti-politician (*r* = .05) statements; I also did the same for the anti-politician prejudice correlation with the same two perceived authenticity scales (*r* = -.06 and *r* = .18, respectively). Confidence intervals for these differences in correlations were calculated using Zou’s (2007) method (Diedenhofen & Musch, 2015). The relationship between Muslim prejudice and thinking anti-Muslim statements are authentic was stronger than the relationship between prejudice against Muslims and perceived authenticity of anti-politician statements, *rdiff* =.33 [.16, .50]. This was also the case with anti-politician prejudice and perceived authenticity of anti-Muslim and anti-politician statements, *rdiff* = .24 [.06, .41].

**Discussion**

Prejudice against Muslims positively predicted *only* perceived authenticity of anti-Muslim statements; the same was true for the equivalent items about politicians. This relationship did *not* hold between reported dislike of control targets (i.e., negative, non-normative statements that lacked the moral significance of prejudice) and perceived authenticity of negative statements about those targets. These results support the first two hypotheses: Prejudice positively predicts perceived authenticity of similarly-prejudiced statements.The data also suggest that there is something particular to prejudice that leads to this relationship, as it was not present for control targets.

**Study 2**

I conceptually generalized the findings from Study 1 by sampling a different population (campus pedestrians), measuring attitudes toward different target groups (students at a rival school, illegal immigrants), and employing a between-subject design. I again predicted that self-reported prejudice would positively predict the perceived authenticity of prejudiced statements—but that this effect would only be present with measures of the same target group.

**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.

**Discussion**

The more someone held a prejudice against a group, the more they perceived negative statements about that group to be authentic. Perceived authenticity of anti-illegal immigrant sentiments was only predicted by prejudice against illegal immigrants; perceived authenticity of prejudice against Kansas State students was only predicted by negative feelings toward Kansas State students. Studies 1 and 2 provide support for my first two hypotheses, suggesting that prejudice leads people to see others expressing prejudices they share to be authentic. Studies 3 – 8 continue to demonstrate this relationship but also extend these findings by examining potential contributing mechanisms for *why* this relationship between prejudice and perceived authenticity exists.

**Study 3**

I investigated the prescriptive norm (hypothesis three) and social projection (hypothesis four) explanations for the relationship between prejudice and perceived authenticity. Recall that my third hypothesis argues prejudice predicts *greater* perceived descriptive (i.e., what participants think people *actually* feel) normativity, which in turn leads to *greater* perceived authenticity. This is due to a perceptual bias: The more people think people have the prejudice, the greater the chances that any one person has prejudice, making expressions of prejudice more likely to be labelled as authentic.

My fourth hypothesis argues that the relationship between prejudice and perceived authenticity gets *more positive* as prejudices are less prescriptively (i.e., what participants think the rules are about how people *should* feel) normative. A second way to test this is by examining how *threatened* people feel by the prescriptive norms. If perceived authenticity operates as a vicarious justification, then the prejudice and perceived authenticity relationship should also be *greater* when people feel this threat (i.e., when participants think there are “politically correct” crusaders in society, unjustly cracking down on speech).

**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. This relationship held even after controlling for prejudice, *b* = .08, *SE* = .02, *t*(128.63) = 3.67, *p* < .001.

**Moderator analyses.** I tested if the relationship between prejudice and perceived authenticity was moderated by (a) perceived prescriptive normativity or (b) concern about political correctness. 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. However, the reader may note that some target groups are not usually discussed in the conversation about “political correctness” (e.g., business people, deaf people); for each of the ten target groups separately, I regressed perceived authenticity onto prejudice, concern about political correctness, and the interaction between the two. The interaction was “significant” for only one of these target groups: lawyers, *b* = -.15, *SE* = .07, *t*(196) = -2.11, *p* = .036. Given the lack of a theoretical prediction for this interaction and performing multiple tests, I consider this to be a Type I error.

**Discussion**

Prejudice predicted greater perceived authenticity of people expressing that same prejudice, again providing support for my first hypothesis. The social projection account received support, as well: The more participants reported a prejudice, the more they perceived others to share that same prejudice; in turn, this perception of descriptive normativity predicted participants perceiving the expression of prejudice as more authentic. A notable shortcoming of this cross-sectional approach, however, is that no compelling causal claims can be made between descriptive normativity and perceived authenticity. This limitation is addressed in the subsequent study, Study 4, by manipulating descriptive normativity.

The data also supported the prescriptive norms account: The less prescriptively normative one perceived the prejudice to be, the greater was the relationship between prejudice and perceived authenticity. However, a complication arises in that this did *not* appear to be driven by how threatened participants felt by the prescriptive norms, since the prejudice and perceived authenticity relationship did *not* depend on concern about political correctness. Studies 5 and 6 aim to clarify this relationship by manipulating prescriptive normativity directly.

**Study 4**

In the present study, I addressed the Study 3 limitations that prevent one from drawing causal support for the social projection account (i.e., hypothesis three). I experimentally tested this hypothesis by manipulating perceived descriptive norms and measuring their effect on perceived authenticity of prejudiced statements. In line with this cognitive account, I predicted that participants would find the prejudiced statements to be *more* authentic when the prejudice was portrayed as descriptively normative than when portrayed as descriptively *non*-normative.

**Method**

I recruited 210 participants from MTurk to participate in a “survey on perceiving other peoples’ attitudes.” The specific prejudice I consider in this study is that against transgender people. To determine the needed sample size, I calculated the zero-order correlation from Study 3 between perceived descriptive normativity of prejudice against transgender people and perceived authenticity of a prejudiced statement toward transgender people, *r* = .22. The equivalent Cohen’s *d* = .45. I recruited enough participants to achieve 90% power at this effect size. A total of 211 people participated. Participants’ ages ranged from 18 to 71 (*M* = 34.41, SD = 11.21), 38% identified as male, and 73% identified as White.

Participants were randomly assigned to either a *high* prejudicenorm or *low* prejudicenorm condition. They were given a brief description containing “results of some previous studies we have done on MTurk,” which read:

Our lab mainly studies prejudice, and we have done many surveys on MTurk about prejudices against various social groups. One of the groups that people generally express the [highest/lowest] prejudice toward in these studies is transgender people. We have seen that MTurkers tend to express pretty [negative/positive] attitudes toward transgender people. Out of the many prejudices that we study, prejudice against transgender people is the one that is expressed [most/least] by people.

After reading this, participants were asked: “If you had to guess, what percentage of MTurkers do you think are prejudiced against transgender people?” They responded on a sliding scale ranging 0% to 100%. I chose “MTurkers” to be the relevant group—not Americans as a whole—because I reasoned that people might easily self-categorize (Hornsey, 2008 ; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) as an “MTurker” while participating in research on the website; thus, the descriptive norm is a group that is relevant to them in the experimental setting.

Participants were then told: “We asked previous MTurkers that took our surveys to explain some of their feelings toward transgender people. We randomly selected one of those quotes.” This quote was: “It’s not a good idea to let transgender people around kids… It may not be safe, and kids will be confused.” Participants indicated perceived authenticity on the same scale as in Studies 1 – 3 and then answered a demographics questionnaire.

**Results**

Participants in the high norm condition (*M* = 54.65, *SD* = 20.50) perceived the descriptive normativity of anti-transgender prejudice to be higher than those in the low norm condition (*M* = 27.38, *SD* = 21.31), *t*(209) = 9.47, *p* < .001, *d* = 1.30 [1.01, 1.60]; the manipulation accomplished what it was intended to do. However, the high norm condition (*M* = 5.25, *SD* = 1.39) did *not* lead to significantly higher perceptions of authenticity (*M* = 5.01, *SD* = 1.46), *t*(209) = 1.23, *p* = .221, *d* = .17 [-.10, .44]. The measure of perceived manipulation did not correlate with perceived authenticity, either, *r* = .08, *p* = .227.

**Discussion**

Descriptive noms did not affect perceived authenticity. While a correlational relationship was found in Study 3, this relationship failed to replicate—both causally and correlationally—in Study 4. The cognitive, social projection account for the prejudice and perceived authenticity relationship was not supported by the data; although prejudiced people see more prejudice in the world, this cognitive bias cannot explain the current phenomenon. I turned to the influence of prescriptive norms and vicarious suppression in the subsequent two studies.

**Studies 5 and 6**

In Study 5, I experimentally tested the prescriptive norms account for the relationship between prejudice and perceived authenticity by manipulating whether or not it was acceptable to express a specific prejudice. I predicted that there would *only* be a relationship between prejudice and perceived authenticity when the prejudice was portrayed as *prescriptively non-normative*. Telling participants it is okay to express a prejudice eliminates the feeling of vicarious suppression and thus should minimize the relationship between prejudice and perceived authenticity.

Study 6 was a close replication of Study 5. I also included measures of perceived political correctness of the expressed prejudice, reasoning that these could be seen as residing at opposite ends of a spectrum: One could be perceived as cowing to the forces of political correctness, allowing themselves to say what they authentically believe, or somewhere in between these two poles. I present both studies separately and then discuss them together.

**Study 5 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 (see also White & Crandall, 2017). 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 – 4 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*).

**Study 5 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 6 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 5, and demographics and prejudice were measured the same as in Study 5. 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.

**Study 6 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 5: 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 5, *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.

**Studies 5 and 6 Results, Synthesized**

In Study 5, framing the expression of a prejudiced statement as acceptable eliminated the established relationship between prejudice and perceived authenticity. However, this interaction was not significant in Study 6. One might ask, “Did this study “fail” to replicate?” One way to answer this question is to simply look at significance: Was the replication also significant and in the same direction? A second way to answer this question is by examining the 95% confidence intervals for the effects of interest: Did the replication’s confidence interval contain the original study’s observed effect? Since all measures were on the same scale (condition was coded 0 or 1, all measures were 7-point Likert scales in both studies), the coefficients for the interactions can be directly compared across studies. The observed effect in Study 5 was *b* = .36 [.14, .59]; the confidence interval for the observed interaction in Study 6 barely included the point estimate from Study 5, *b* = .14 [-.10, .37]. By both criteria, Study 6 failed to replicate Study 5.

But asking “Did it replicate?” and expecting a simple “Yes” or “No” answer is misguided. Differences between an original and replication study could occur for a wide variety of reasons—such as lack of power or sampling and measurement error—*even when* both studies are capturing the same population effect size (e.g., Maxwell, Lau, & Howard, 2015; Stanley & Spence, 2014). This leads to judgments about “success” or “failure” of replication, particularly in the case of two studies, to be incredibly vague.

Instead, I ask: “Is there cumulative evidence for the hypothesis?” (e.g., Braver, Thoemmes, & Rosenthal, 2014; Schmidt, 1996). Answering this question relies on synthesizing the data, not making a series of significant-or-not judgments. Neither the original nor the replication should be privileged—their evidential value should be taken into consideration together. Given the similarity by which the studies were conducted and the difficulties in meta-analyzing regression slopes (e.g., Becker & Wu, 2007), I simply combined the data from Studies 5 and 6, then I regressed perceived authenticity on condition, self-reported prejudice, and the interaction between the two. This interaction was significant, *b* = .24, *SE* = .08, *t*(395) = 2.87, *p* = .004. Probing this interaction revealed that prejudice predicted perceived authenticity in the suppression condition, *b* = .31, *SE* = .06, *t*(395) = 5.20, *p* < .001, but not in the expression condition, *b* = .07, *SE* = .06, *t*(395) = 1.16, *p* = .245. The effect is present, yet the effect is smaller in magnitude than was planned for a priori.

I also regressed perceived authenticity on condition, prejudice, and which study the data came from (i.e., Study 5 or Study 6), as well as all possible two- and three-way interactions. The condition by prejudice interaction remained significant, *b* = .36, *SE* = .12, *t*(391) = 3.10, *p* = .002, and was not qualified by a three-way interaction with study, *b* = -.23, *SE* = .17, *t*(391) = -1.38, *p* = .170.

**Studies 5 and 6 Discussion**

Framing the expression of a prejudiced statement as acceptable eliminated the established relationship between prejudice and perceived authenticity. This suggests that labelling expressed prejudice as authentic is motivated by vicarious justification—people high in prejudice push back against the prescriptive norms, even when it concerns someone else’s speech act. Contrary to predictions, perceived political correctness was unrelated to perceived authenticity. While the more prejudiced people were, the more they saw the expression as politically correct, this relationship was unaffected by presenting prejudice as acceptable or unacceptable to express.I turn to testing two additional intuitive explanations for the prejudice and perceived authenticity relationship, balance and motivated reasoning, are tested in Studies 7 and 8, respectively.

**Study 7**

I experimentally tested the balance account (hypothesis five)—that the positive relationship between prejudice and perceived authenticity is due to the need for affective consistency. If someone is prejudiced, they have a positive relationship with a prejudiced statement (by virtue of agreeing with it); if authenticity is presented as a *positive* thing, people maintain affective balance by claiming a positive relationship between the prejudiced statement and authenticity. I predict that the established relationship between prejudice and perceived authenticity should only be present when authenticity is portrayed as a positive quality, but *not* when portrayed negatively.

**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.

**Discussion**

Prejudice predicted perceived authenticity, regardless of authenticity’s affective valence. Contrary to the predictions following from balance theory, the relationship between prejudice and perceived authenticity remained positive even when authenticity was presented as a negative quality. I again demonstrated the relationship between perceived authenticity and prejudice was demonstrated; however, it cannot be explained by the intuitive explanation of, “People attach positive labels to the things with which they agree.” In the final study, I examine the motivated reasoning account for the present phenomenon.

**Study 8**

I examined the motivated reasoning (hypothesis six) account by manipulating participants’ goals in judging the statement for its authenticity. If people’s perceptions of authenticity are motivated by their need to express prejudice, then directing people toward a different perceptual goal should weaken the positive relationship between prejudice and perceived authenticity of prejudice expression. I predicted that incentivizing accuracy, in addition to providing extra information for participants to reason over, would lead to a smaller prejudice and perceived authenticity relationship than incentivizing participants to express themselves. Much of the research reviewed by Kunda (1990) encouraged accuracy goals indirectly—by telling participants they would have to defend their judgments to others, that their judgments would be evaluated or made public. As a more direct test of the theoretical argument, I directly encouraged accuracy by explicitly asking participants to focus on accuracy.

**Method**

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 – 5. I timed how long participants spent on this page, since deeper processing for accuracy goals is hypothesized to take longer (Kunda, 1990; but see Mullen & Skitka, 2006).

Participants then 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**

Participants in the accuracy goal condition (*M* = 96.26, *SD* = 69.06) did not spend more seconds reading the passage and answering the authenticity questions than those in the expressive goal condition (*M* = 92.02, *SD* = 80.59), *t*(218) = 0.42, *p* = .675, *d* = .06 [-21, .32]. However, it is unclear how this should be interpreted (see Discussion below).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. There was again a zero-order relationship between prejudice and perceived authenticity, *r* = .20, *p* = .003.

**Discussion**

Manipulating perceiver’s goals when judging authenticity did not affect the relationship between prejudice and perceived authenticity.These data do not provide support for the motivated reasoning account—that prejudiced people engage in biased processing because they wish to perceive authenticity—for the core phenomenon of interest. However, it remains ambiguous as to if goals were successfully manipulated. A priori, I assumed that people would take longer reading, considering, and processing the speaker’s biographical information in the accuracy condition due to spending more cognitive resources on doing so (as is implied by Kunda, 1990, pp. 481-482). There was no significant difference in time spent reading between conditions in the present study. However, Mullen and Skitka (2006) argued that *longer* time spent processing as evidence for motivated reasoning. Their research focused on Dr. Jack Kevorkian’s guilty verdict of second-degree murder for his role in physician-assisted suicide. They hypothesized that those who were in support of physician-assisted suicide would have a directional goal to find flaws in the case, and thus would spend *more time* picking apart information from the case. Given that the current study used a biographical depiction of the speaker that was designed to be *neutral* in an attempt to not cause any ceiling or floor effects on the judgments of perceived authenticity (i.e., not to make it too obvious whether or not the speaker was actually prejudiced), it remains a possibility that motivated reasoning could play a role *only* when there is specific information that points to or against authenticity. However, this line of reasoning *cannot* explain why people who are told nothing else about the speaker assume that the expressions of prejudice are authentic.

**Meta-Analysis**

I performed a meta-analysis (N = 1386) on the primary correlations of interest—those between a self-reported prejudice and perceived authenticity of that same prejudice—to assess the strength of the relationship between the two. I used a multivariate approach (Jackson, Riley, & White, 2011; Konstantopoulos, 2011, Viechtbauer, 2010) to account for the dependencies between correlations in Studies 1 and 3. Each correlation coefficient was transformed to Fisher’s *z*, and variances and covariances for effect sizes were calculated according to Olkin and Finn (1990). Eighteen correlations were analyzed (Table 2). The meta-analytic correlation was *r* = .22 [.18, .26]. Given that a number of experimental conditions in the paper aimed to eliminate the relationship between prejudice and perceived authenticity, however, this point estimate might be an underestimate.

**General Discussion**

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.

Table 2

*Correlations and 95% Confidence Intervals of Meta-Analyzed Correlations*

|  |  |  |  |
| --- | --- | --- | --- |
| Study | Target | *r* | 95% CI |
| 1 | Muslims | .38 | [.22, .52] |
|  | Politicians | .18 | [.00, .34] |
| 2 | Illegal immigrants | .36 | [.18, .51] |
|  | Kansas State students | .42 | [.25, .57] |
| 3 | Black people | .15 | [.01, .28] |
|  | Transgender people | .19 | [.05, .32] |
|  | Fat people | .05 | [-.09, .19] |
|  | Police officers | .24 | [.11, .37] |
|  | Lawyers | .22 | [.08, .35] |
|  | Business people | .16 | [.02, .29] |
|  | Prostitutes | .21 | [.08, .34] |
|  | Drug dealers | .13 | [-.01, .13] |
|  | Blind people | .07 | [-.07, .21] |
|  | Deaf people | .04 | [-.10, .18] |
| 5 | Fat people | .24 | [.11, .37] |
| 6 | Fat people | .25 | [.11, .37] |
| 7 | Illegal immigrants | .21 | [.08, .33] |
| 8 | Black people | .20 | [.07, .32] |
| Meta-analysis |  | .22 | [.18, .26] |