**Perceived Authenticity as a Vicarious Justification for Prejudice**

Authenticity is prized in American culture. We praise those who resist social pressures, remain honest to their convictions, and live in accordance with their inner, core selves. But what of those whose innermost attitudes, such as prejudice, are seen by many as offensive? Appeals to authenticity in these cases seem dissonant with the positive connotation authenticity carries with it—yet it seems that people see others expressing prejudice as being authentic, genuine, honest, and true to themselves.

The 2016 United States presidential election provides an example. Donald Trump expressed prejudice throughout his campaign—and continues to into his presidency. He has openly insulted Mexicans, immigrants, Muslims, Haitians, Africans, Black Americans, Puerto Ricans, Korean Americans, women, people who are overweight, among others (Lee & Quealy, 2018; Leonhardt & Philbrick, 2018). Polling suggests many Republican primary voters thought Trump was “authentic” (Sargent, 2015), and people across the political spectrum—conservative to progressive—called Trump “*the* authentic” candidate during election season (Estepa, 2017; Johnson, 2017; Merelli, 2017). Trump’s flouting of norms against explicitly expressing prejudice has been one of the oft-mentioned reasons for labeling him “authentic” (Basavaraju, 2016; Bebout & Garcia, 2016; Gallagher, 2017; Johnson, 2016). Trump qualifies offensive statements by saying that he is just being “honest” or “frank” (Reston, 2015). Many others deploy this tactic, such as: journalists claiming that anti-prejudice norms prevent an “honest” conversation about prejudice (Abernathy, 2017; Dickerson, 2004), reality television show contestants expressing racial stereotypes (Fallon, Gray, & Offenberg, 2017), and White nationalists discussing their racial ideology (Newton, 2017; Taylor & Molyneux, 2016).

Why do people perceive another’s expression of prejudice to be authentic? I propose people will do so as an attempt to vicariouslyjustify another’s expression of 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 situating this concept within modern theories of prejudice expression. I then review four psychological theories that may explain *why* people might variously justify another’s expression of prejudice by labeling it “authentic.”

**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 (see Kernis & Goldman, 2004). In this tradition, authenticity was broadly defined as living in accordance with one’s true, core, inner self. These psychologists saw the inner core of a person as being fundamentally good, and living in accordance with this core being an instinct people should foster in order to thrive.

Many ideas in the humanist tradition were treated synonymously with authenticity. Carl Rogers proposed the concept of *congruence*, which refers to when one’s ideal version of themselves aligns with how one sees themselves. Rogers (1961) described the state of congruence by quoting Kierkegaard: “to be that self which one really is.” Although he did not explicitly call this authenticity, it matches contemporary definitions of the term, and much of his work describes how closely-related concepts like being genuine, true, real, and not “putting up facades” to others leads to well-being (Rogers, 1961).

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 (1964; 1971; 1974) defined authenticity as 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 drivers of happiness is living an authentic life—for people to let others see them how they see themselves and to resist the external pressures to suppress their true self.

These humanistic ideas about authenticity were not subjected to many empirical tests, 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); to these researchers, authenticity is 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**

The present research question—what drives people to label others’ expressions of prejudice as authentic?—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 found across various fields, including psychology, sociology, marketing, and communications.

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; Grayson & Martinec, 2004). Qualitative research shows the breadth of ways people understand authenticity.

Beverland and Farrelly (2010) asked participants how they defined authenticity, presenting them with 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 *in*authenticity 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)—yet people are very willing to assign authenticity to museum 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 faithfully 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.

Iconic authenticity was exemplified by museum patrons 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” (Grayson & Martinec, 2004, pp. 300-301). Patrons also noted the authenticity of the Holmes museum because everything inside looked as if it belonged at that historical time; 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. Amateurism and non-professionalism—antonyms to these words found by Kovacs and colleages—can lend an aura of authenticity to a political candidate (Enli, 2017; Manning, Penfold-Mounce, Loader, Vromen, & Xenos, 2017). The meaning of authenticity is nebulous, fluid, and it can be assigned to many things that are simply “good,” depending on the context.

Researchers across disciplines have also studied the correlates, causes, and consequences of perceived authenticity. 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 lower in 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).

Hahl, Kim & Sivan (2017) examined how people could find a “lying demagogue” to be 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 was led to feel as if 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 disregard the established norms of a political system that participants found illegitimate.

Pillow, Crabtree, Galvin, and Hale (2017) argued that people are motivated to see political candidates 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, interpreting the interaction to be due to a cognitive process whereby people who like the candidate weigh perceived candor as more important 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 perceived authenticity 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 to perceptions of authenticity.

**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 (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.

*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 (Choi, Crandall, & La, 2014; Miller & Effron, 2010; Monin & Miller, 2001).

**Authenticity as Justification for Prejudice**

I propose that prejudiced people are more likely to perceive an expression of prejudice as “authentic” because perceived authenticity can be a justification for prejudice. The vague nature of how authenticity is defined allows for flexibility in how it is interpreted, making it amenable to justifying prejudice. The concept is a positive, abstract concept that can be appropriated for a prejudiced person’s need to defend expressions of prejudice. 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). Members of the president’s party struggled to defend his racist remarks when he wondered aloud in a meeting with lawmakers why the United States was welcoming Haitians and Africans—in his reported words, people from “shithole countries.” Some tried to reframe the statement as merely being “tough language,” while others tried to cast it as not racial but an argument for merit-based immigration policies (Fang, 2018; Phillips, 2018; Watkins & Philip, 2018).

White and Crandall (2017) conducted eight experiments where 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; instead, they defended 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 “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. They focus less on whether 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, “N\*\*\*\*r, n\*\*\*\*r, n\*\*\*\*r.” By 1968 you can’t say “n\*\*\*\*r”—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 “N\*\*\*\*r, n\*\*\*\*r” (in Perlstein, 2012).

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 because it allows someone to express that they agree with the prejudiced statement in an ostensibly race-neutral way—without ever explicitly mentioning the target group—that is less likely to bring about social opprobrium or feelings of guilt.

**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: social projection, prescriptive norms, balance, and motivated reasoning.

**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;. 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 justification-suppression model, this social projection explanation for the positive relationship between prejudice and perceived authenticity is an unmotivated, cognitive one—it is a simple perceptual bias toward perceiving others as 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 then *increase* perceived authenticity of prejudiced statements.

**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). Following from the justification-suppression model, *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.”

**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.” 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).

I propose that a prejudiced person perceives authenticity in a prejudiced statement in order to maintain balance. In Heider’s notation: 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 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, where the information people used to make their judgments of authenticity depended on their pre-existing attitudes. Their test was correlational; I will test this account experimentally.

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 justification-suppression model, however, the goal in the present situation is to *express* prejudice. Crandall and Eshelman (2003) directly argued 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.

**Dispositional Attributions**

Attributional theories suggest another account for why people perceive prejudiced statements to be authentic: Perceived authenticity may be the result of an observer making an internal, dispositional attribution. In contrast to the primary accounts proposed above, the perceiver’s own prejudice does *not* drive this perception here. Dispositional attributions are made when observers perceive the cause of a target’s actions to be coming from *within* the target, from their personality. This definition is similar to that of perceived authenticty, meaning that calling an target “authentic” might be a colloquial way of communicating one has made a dispositional attribution.

People tend to perceive actions as coming from either external factors (e.g., norms, rules, others’ behavior) or internal factors (e.g., personality, desires, goals of an actor). Norms are an *external* force leading people to act in a certain way; if a target acts *opposite* this force, others are likely to make an internal, dispositional attribution for this action, since the external force cannot account for it (Heider, 1958, Chapter 4; Jones, 1976; Jones & Davis, 1965; Jones & Harris, 1967; Kelley, 1973). When a target expresses prejudice that is either prescriptively or descriptively non-normative, they are acting in a way that cannot be explained by situational factors or the behavior of others. People are likely to see this expression as coming from within the target—a dispositional attribution, perceiving authenticity. This implies that perceived normativity of a given prejudice should be *negatively* correlated with the perceived authenticity of a target expressing prejudice.

This prediction is opposite to that of the social projection account—but it is in line with Pillow et al.’s (2017) findings, which they frame within correspondent inference theory (one of the many attributional approaches that would predict this relationship).

Dispositional attributions should operate orthogonally to one’s own prejudice in their influence on perceiving prejudiced statements as authentic—both attributions and prejudices could operate simultaneously. This account is thus not an *alternative* for the accounts proposed above, but another intuitive reason for perceiving authenticity. It will be empirically tested alongside the social projection account.

**The Present Studies**

I test six hypotheses across eight studies. These hypotheses—and their associated theoretical perspectives—are presented in Table 1. The first two hypotheses test the general phenomenon that perceived authenticity may justify prejudice, while the latter four examine psychological processes that may underlie the proposed positive relationship between prejudice and perceived authenticity. I refer to these hypothesis as H1 through H6 throughout the paper.

**Study 1**

I tested H1 and H2 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, trivial tastes.

**Results**

Correlations that test the primary hypothesis are reported in the diagonal of Table 2, which contains 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 (Figure 1). 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 significantly above zero, which would be the case if this general tendency was present. This does not appear to be an issue of statistical power, either: All the positive off-diagonal correlations were smaller than *r* = .05. As a formal statistical test of H2, 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].

It is not clear why both anti-Muslim and anti-politician prejudice correlated *negatively* with perceived authenticity of negative statements about the beach. Since this does not challenge the current hypotheses being investigated, I do not consider it further.

**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 H1 and H2: Prejudice positively predicts only the 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 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 (H1)—but that this effect would only be present with measures of the same target group (H2).

**Method**

Participants were recruited as they were walking around the University of Kansas campus. Research assistants approached passersby 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 could be recruited by the end of the semester. A total of 221 people participated, but 7 participants were excluded for partial nonresponse.

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. I tested these predictions across two separate regression equations, each using a different type of prejudice (illegal immigrant or Kansas State) as the focal predictor. 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 (Figure 2).

**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 H1 and H2, 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 social projection (H3) and prescriptive norm (H4) explanations for the relationship between prejudice and perceived authenticity. H3 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 others in general have the prejudice, the greater the chances that any one person has prejudice, making expressions of prejudice more likely to be labelled as authentic. H4 argues that the relationship between prejudice and perceived authenticity is *more* positive as prejudices are less prescriptively (i.e., what participants think the rules are about how people *should* feel) normative, as proscription of prejudice creates the motive for justification.

**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 on MTurk 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 subjectively. I recruited 200 participants and each participant contributed 10 data points; this generates a level one *n* = 2000 and a level two *n* = 200. 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.** Questions about norms, authenticity, and prejudice 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 on a sliding 0 to 100 scale “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.

**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 authenticity 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.

**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 (acrossindividuals) before analyses. Ten measurements—one for each target group—were nested within each individual. Thus, perceived descriptive and prescriptive normativity, perceived authenticity, and prejudice were modeled at the “first” level, with participants being the “second” level. All regression coefficients, as well as the intercept, were allowed to differ by individual; that is, random intercepts and random slopes for all level-one predictors were defined. For example, every participant had their own intercept and slope estimated for the relationship between prejudice and perceived authenticity. I used Satterthwaite’s approximation for degrees of freedom for all *t*-tests of regression coefficients (Kuznetsova, Brockhoff, & Christensen, 2017).

**Results**

**Social projection.** Supporting H1, prejudice again correlated positively with perceived authenticity of negative statements, *b* = .24, *SE* = .03, *t*(181.53) = 8.31, *p* < .001. This was the average coefficient across participants (i.e., the fixed effect). There was significant variance among participants about this coefficient (Figure 3).

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. In turn, both prejudice, *b* = .20, *SE* = .03, *t*(162.43) = 6.60, *p* < .001, and descriptive normativity, *b* = .08, *SE* = .02, *t*(128.63) = 3.67, *p* < .001, were significant predictors on perceived authenticity when considered in a regression equation simultaneously (these are again the average coefficients across participants, as are the rest of the coefficients reported in this section). This chain of relationships supports H3: Prejudice positively predicts perceived descriptive normativity, which in turn positively predicts to perceived authenticity.

**Prescriptive norms.** I tested if the relationship between prejudice and perceived authenticity was moderated by perceived prescriptive normativity. I regressed perceived authenticity on prejudice, prescriptive normativity, and the interaction between the two. All coefficients were allowed to vary by individual (i.e., each were estimated with random effects).

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 (Figure 4). This pattern of simple slopes supports H4: Perceptions of authenticity are strongest when the expression of prejudice is counter-normative.

**Discussion**

Prejudice predicted greater perceived authenticity of people expressing that same prejudice, again providing support for H1. The social projection account (H3) received suggestive 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 (e.g., Bullock, Green, & Ha, 2010; Holland, 1988; Imai, Keele, & Tingley, 2010; Ten Have & Joffe, 2010). This limitation is addressed in the subsequent study, Study 4, by manipulating descriptive normativity.

Attributional theories (e.g., balance, Heider, 1958; correspondent inference, Jones & Davis, 1965) predict the opposite relationship: Perceived authenticity arises from making dispositional attributions, which arise when one perceives an actor to be going *against* norms. The dispositional attribution account was thus not only *un*supported—but evidence was found to the contrary of its predictions.

The data also supported the prescriptive norms (H4) account: The less prescriptively normative one perceived the prejudice to be, the greater was the relationship between prejudice and perceived authenticity. Studies 5 and 6 aim to experimentally test this account by manipulating prescriptive normativity directly.

**Study 4**

I addressed the Study 3 limitations that prevent one from drawing causal support for the social projection account (H3). I experimentally tested H3 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 considered in this study is that against transgender people; to determine the needed sample size, I calculated the zero-order correlation in 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, and 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 descriptive norm condition (*M* = 54.65, *SD* = 20.50) believed a larger percentage of MTurkers to be prejudiced against transgender people 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, participants in the high descriptive norm condition (*M* = 5.25, *SD* = 1.39) did *not* perceive the person expressing anti-transgender sentiment to be more authentic than participants in the low norm condition (*M* = 5.01, *SD* = 1.46), *t*(209) = 1.23, *p* = .221, *d* = .17 [-.10, .44] (Figure 5). The measure of perceived descriptive normativity (the manipulation check) did not correlate with perceived authenticity, *r* = .08, *p* = .227.

**Discussion**

Descriptive norms did not affect perceived authenticity. While a correlational relationship was found in Study 3, this relationship was not supported in an experimental setting in Study 4. The cognitive, social projection account (H3) 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 is not a compelling explanation for the current phenomenon—H3 did not receive causal support from the randomized experiment. Additionally, the dispositional attribution account was again not supported by the data, given that a negative correlation between perceived normativity and perceived authenticity was not present. I now turn 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 (H4) for the relationship between prejudice and perceived authenticity by manipulating whether or not it was acceptable to express a specific prejudice. In line with H4, 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 should thus 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 expressed prejudice. The ostensible authenticity of prejudiced speech has often been pitted against the supposed disingenuous, overly-polite speech influenced by political correctness (e.g., Tumulty & Johnson, 2016). I hypothesized that authenticity and perceived authenticity 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 included perceived political correctness as another dependent variable, predicting it to behave in a fashion opposite that of perceived authenticity. 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* = 6.10, *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 et al., 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. Prejudice was only a predictor of perceived authenticity when the prejudiced statement broke the rules—when it was prescriptively non-normative.

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 nonresponse. 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 (H4) that the relationship between prejudice and perceived authenticity would only hold when the prejudice was presented as prescriptively non-normative 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, supporting H4. However, this interaction was not significant in Study 6, failing to support H4. 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? By this criteria, it was a “failure” to replicate. 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 this second criteria, it was a “successful” replication, albeit not a compelling one.

But asking “Did it replicate?” might not be the most valid approach to determine if H4 was generally supported. 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 *exact* same population effect size (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?” (Braver, Thoemmes, & Rosenthal, 2014; Schmidt, 1996). Answering this question relies on synthesizing the data, not tallying a series of significant-or-not judgments. Neither the original nor the replication should be privileged in this case—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 (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 (Figure 6). 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 smaller in magnitude than was planned for a priori.

I ran two more regression analyses to examine the role of study. Simply controlling for study by adding it as a dichotomous covariate to the above model left the two-way interaction between condition and prejudice unchanged at the precision of two decimal points, *b* = .24, *SE* = .08, *t*(394) = 2.87, *p* = .004. 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 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 prescriptive norms and vicarious justification—people high in prejudice push back against the prescriptive norms, even when it concerns someone else’s speech act (H4). 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; the balance (H5) and motivated reasoning (H6) accounts are tested in Studies 7 and 8, respectively.

**Study 7**

I experimentally tested the balance account (H5), which argues 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 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 about illegal immigrants that were 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]. Despite the condition having a sizable effect on this manipulation check, the mean was still well above the middle of the scale—and only 11 participants (9.6%) in the *authenticity is bad* condition responded below the midpoint. The condition reduced the goodness somewhat, but authenticity was still seen as good. Participants 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).

I tested H5 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 (Figure 7). However, prejudice again positively predicted perceived authenticity; the zero-order correlation between the two (i.e., collapsing across condition) was *r* = .21, *p* = .002.

**Discussion**

Prejudice predicted perceived authenticity, regardless of authenticity’s affective valence. Contrary to the predictions following from balance theory (H5), the relationship between prejudice and perceived authenticity remained positive even when authenticity was presented as a negative quality. However, very few participants were able to be persuaded that authenticity was a negative quality—balance theory may still be supported if participants were convinced that authenticity was not merely “less good” but “negative.”

I again demonstrated the relationship between perceived authenticity and prejudice (H1). However, the data did not support an intuitive explanation—that people call prejudices they agree with authentic simply because they think authenticity is a “good thing”—for this relationship. In the final study, I examine the motivated reasoning account for the present phenomenon.

**Study 8**

I examined the motivated reasoning (H6) 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 about, 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 explicitly asked 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 (Yzerbyt, Schadron, Leyens, & Rocher, 1994; Leyens, Yzerbyt, & Schadron, 1991). 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 implied to take longer (Kunda, 1990; c.f. Mullen & Skitka, 2006).

Participants answered a demographics questionnaire. Prejudice was then measured using an eight-item, seven-point symbolic racism scale (Henry & Sears, 2002; e.g., “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].

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 (Figure 8). 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 (H6). 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 cannotspeak to 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 3). The meta-analytic correlation was *r* = .22 [.18, .26], supporting H1. 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**

Perceiving an expression of prejudice as authentic is partially driven by the perceiver’s own prejudice. The more prejudice participants reported, the more they saw an expression of prejudice as authentic; I demonstrated this across multiple samples (N = 1386), a variety of target groups, and various measures of prejudice. This was not due to prejudiced people finding negative statements in generalas authentic: The positive relationship between prejudice and perceived authenticity was only found when the target groups for both measures were the *same*. This relationship was only found among prejudiced attitudes, as well: Negative, descriptively non-normative statements about trivial targets (i.e., the beach, cookies, ice cream) were *not* predicted by how much one disliked those same targets. Why is this relationship present for prejudice but not other attitudes? Why do people label others’ expressions of shared prejudices as authentic? I tested four different accounts across six different studies to answer these questions.

**Prescriptive Norms: Authenticity as a Justification for Prejudice**

Only the prescriptive norms account yielded reliable empirical support. The justification-suppression model (Crandall & Eshleman, 2003) posits that many prejudices face suppressive forces that prevent people from expressing them. But this does not necessarily eliminate prejudice or its motivation to be expressed. People thus find justifications for these negative attitudes in an attempt to express them without facing punishment from others and free from feelings of guilt. If a prejudice is socially unacceptable (i.e., prescriptively non-normative) to express, then people seek out justifications. If a prejudice *is* socially acceptable, justifications are not necessary—people can directly express their prejudice (Crandall et al., 2013). Studies 5 and 6 together demonstrated that perceptions of authenticity behave how justifications for prejudice behave: The positive relationship between prejudice and perceived authenticity was *only* present when prescriptive norms against expressing that prejudice were present. When I told participants that it was acceptable to blame overweight people for their own weight problems, authenticity no longer correlated with self-reported prejudice. However, when I told participants that it was *not* acceptable—that, in fact, it was prejudiced—to blame overweight people for their weight, then prejudice predicted perceived authenticity. Study 3 demonstrated this cross-sectionally; the more participants perceived the prejudice to be prescriptively non-normative, the stronger the relationship was between prejudice and perceived authenticity. These data show that perceptions of authenticity behave as do justifications for prejudice, suggesting that people might claim an expression of prejudice as authentic as a way to justify someone else’s prejudice (vicarious justification) or to express prejudice in a coded way (justification by rearticulation).

**Social Projection, Balance, and Motivated Reasoning**

The social projection, balance, and motivated reasoning accounts did not yield empirical support. Social projection (Kreuger, 2007) occurs when people presume others are like themselves. In the present case, prejudiced people are more likely to think people in general are prejudiced. The social projection hypothesis argues that the more prejudiced that one perceives people to be in general, the more likely they believe it is that any one person is prejudiced. This argues prejudiced people see authenticity due to a cognitive bias—they think it is more common, so any expression of prejudice is more likely to be authentic or real. Study 3 results provided suggestive evidence for the social projection account. The more prejudiced participants reported, the more common they thought the prejudice was in society (i.e., descriptive normativity; this evidence supports social projection and directly refutes the dispositional attribution account). Perceived descriptive normativity, in turn, predicted greater perceived authenticity of prejudiced statements. I manipulated descriptive normativity of prejudice in Study 4, however, and it had no effect on perceived authenticity. This cognitive account does not seem to explain the positive relationship between prejudice and perceived authenticity, and perceived authenticity does not appear to be an expression that one has made a dispositional attribution.

Like the prescriptive norms hypothesis, predictions generated from the balance and motivated reasoning accounts both relied on the logic that the correlation between prejudice and perceived authenticity would be eliminated—or at least shrunk—if the underlying causal mechanism was removed. Studies 7 and 8 did not find evidence for either the balance or motivated reasoning accounts. Balance theory (Heider, 1958, Chapter 7) suggests that the need for affective balance simply leads people to label positively things with which they agree—prejudiced people call expressions of prejudice authentic. However, framing authenticity as a negative concept did *not* affect the relationship between prejudice and authenticity. Motivated reasoning (Kunda, 1990) suggests that people will process information in a biased fashion when they are motivated to make a certain inference about someone (e.g., if a target is being authentic), while people process information more dispassionately when they are motivated to be accurate. Prejudiced people might be motivated to see others as authentic, leading to the relationship between prejudice and perceived authenticity. However, motivating participants to be accurate in their judgments—as compared to expressing themselves with their judgments—did *not* affect the relationship between prejudice and authenticity. The social projection, balance, and motivated reasoning accounts were not supported by the data.

**Vicarious Justifications**

White and Crandall (2017) argued that people may justify someone else’s prejudice as a function of their own. When an employee was fired for saying something offensive on social media, anti-Black prejudice predicted more agreement that this firing went against the employee’s right to freedom of speech—but only when the offensive remark was also anti-Black. This could, in part, be explained by prejudiced people feeling a threat to their expressive autonomy: Reading about a person fired for expressing prejudice led similarly-prejudiced people to feel as if they could not express themselves, which in turn predicted people suggesting the firing violated free speech rights.

I consider claiming another’s expression of prejudice to be authentic as a justification for another’s prejudice, much like freedom of speech claims serve the same purpose. White and Crandall (2017) presented participants with a punishment for the prejudiced speaker, making freedom of speech a relevant justification in that circumstance. In the present studies, the only information participants could rely on was the statement itself—and their own notion of what it meant to be authentic. In both sets of studies, however, the forces of prescriptive norms were present: Telling participants that someone was fired for saying anti-Black statements (White & Crandall, 2017) and telling participants that it is against the rules—and prejudiced—to blame overweight people for being overweight (Studies 5 and 6) both communicate a prescriptive norm that it is unacceptable to express prejudice. Anti-Black prejudice and anti-fat prejudice only predicted freedom of speech and perceived authenticity, respectively, when prescriptive norms against expressing those prejudices were made salient.

The present studies, along with White and Crandall (2017), provide converging evidence that people can feel the suppressive forces placed on someone else expressing a shared prejudice. In accordance with the justification-suppression model (Crandall & Eshleman, 2003), people then seek out justifications to circumvent suppression. The content of the justification will change with the context. Freedom of speech is especially relevant in the case of one being fired for expressing prejudice; perceived authenticity might be especially relevant in the case of politicians, who are often labelled as inauthentic, insincere, and Machiavellian (Enli, 2017; Hahl et al., 2017; Manning et al., 2017; Serazio, 2017). The present studies did not present participants with a context, which may be one of the many reasons why the observed correlation between prejudice and the justification is smaller in the present paper (meta-analytic *r* = .22) than in White and Crandall’s (2017) article, (meta-analytic *r* = .43). A ripe area for future research is examining how the content of justifications change with the context in which prejudice is expressed.

**Establishing the underlying cause.** I argue that perceived authenticity, like freedom of speech, is a vicarious justification for another’s expressed prejudice. In doing so, I argue that suppressive forces—particularly prescriptive norms against prejudice—are the *underlying cause* for the relationship between self-reported prejudice and these justifications.Reliably demonstrating meaningful relationships is difficult; determining *why* these relationships occur is even more so, requiring an accumulation of carefully-collected evidence (Bullock et al., 2010).

In Study 3, prejudice was a stronger predictor of perceived authenticity among those who perceived the perceived prejudices to be prescriptively non-normative; in Studies 5 and 6, the relationship between prejudice and perceived authenticity was eliminated when prescriptive non-normativity was absent. White and Crandall (2017, Study 3) also demonstrate this with freedom of speech as a justification, and they demonstrate (Study 8) that it is partially due to prejudiced people being vicariously suppressed (i.e., feeling as if their autonomy is threatened).

Each of these studies have limitations. First, Study 3 is cross-sectional. Second, studies 5 and 6 operate on a causal logic that does not follow from a strict, deductive logic: Prejudice and perceived authenticity correlate when prescriptive non-normativity is present, but they do not when prescriptive non-normativity is absent; therefore, prescriptive non-normativity is the underlying cause for that relationship. Third, White and Crandall (2017, Study 8) employ a moderated mediation model, which suffers from strict statistical assumptions that are assuredly not met (Imai et al., 2010). However, I argue that the present studies are nonetheless consistent with—and support—a vicarious justification account for why perceived authenticity is positively predicted by self-reported prejudice. These studies shape a cumulative body of evidence for the proposed account. Future research should continue to conceptually replicate the present phenomena.

**Lower-prejudiced people.** One might note that Figures 4 and 6 show that lower-prejudiced people are the cases where prescriptive normativity makes a difference. People in my sample who reported the most prejudice tended to find the statements authentic *regardless* of prescriptive normativity. It is clear that lower-prejudiced people recognize and shift *away* from justifications for prejudice (see also White & Crandall, 2017, p. 425). This finding is interesting in its own right, as it suggests that lower-prejudiced people can see through the veneer of justifications to the underlying prejudice; however, it is likely that non-significant differences among highly-prejudiced people are due to the underrepresentation of prejudiced people in the present studies. Both Mechanical Turk and college campus samples tend to be less prejudiced (and less represented in demographic characteristics that correlate positively with prejudice) than the general population (Clifford, Jewell, & Waggoner, 2015; Henry, 2008), and the prejudice measures employed here tended to be positively-skewed (see Figures 6 – 8). Researchers should seek out ways to recruit participants at the high end of the prejudice spectrum, so that phenomena concerning highly-prejudiced people can be better understood.

**Justification by Rearticulation**

Justifications rarely let people openly express prejudice, given the current political climate that sharply punishes explicit bigotry; instead, people use ostensibly non-prejudiced phrases to express their feelings toward social groups. For example, instead of claiming that “Black people are lazy, and that is why they don’t have jobs,” one could say that there is a “tailspin of culture, in our inner cities in particular, of men not working… or learning the value and the culture of work...” (Blow, 2014). The meta-analytic correlation here shows that prejudice predicts something that purports to be prejudice-neutral. In claiming that a speaker was being authentic, one never explicitly *agrees* with the prejudiced statement (even though I find that agreement is one of the predictors of claiming authenticity). In a culture where explicit prejudice is proscribed, it is vital that researchers continue to examine the prejudices underlying supposed prejudice-neutral conversations (White & Baldwin, 2018).

**Portraying Prejudice as Authentic**

The majority of extant work on justifications for prejudice examine why people justify prejudice and the content of justifications for prejudice. But if prejudices are social or psychological processes that allow people to express prejudice without negative sanctioning (Crandall & Eshleman, 2003), an important question is to ask: Do justifications work? For whom do justifications work? How would portraying an expression of prejudice as authentic influence how others perceived the expresser of prejudice? Would it actually prevent the negative sanctioning that the expresser believes the justification will prevent? A fuller understanding of the justification-suppression model would benefit from research focusing on the interplay between expresser and audience: Does the expresser tailor their justifications to the audience? How does the audience react to the justification? This conception of the phenomena could use the justification-suppression model to map both sides of the “national conversation” about prejudice.

**Conclusion**

Authenticity is a nebulous concept. Boyle (2003) claimed that trying to “pin down” what authenticity means is “fiendishly paradoxical” (p. xviii), which inspired him to spend nearly 300 pages exploring what authenticity means in modern society. The present studies showed that one of the many uses of authenticity is to express agreement with and justify the expression of prejudice. The more prejudiced someone was, the more they perceived an expression of that same prejudice as authentic. This was only present when the prejudice was described as socially unacceptable, as prescriptively non-normative. Authenticity, like many other abstract concepts, can carry with it prejudiced implications, and people can use this concept strategically to defend socially unacceptable attitudes that they share.

Table 1

*Hypotheses, Theoretical Perspectives, and Associated Studies*

|  |  |  |  |
| --- | --- | --- | --- |
| Number | Theoretical Perspective | Concrete Hypothesis | Studies |
| 1 | Vicarious justification; justification by rearticulation | Self-reported prejudice should positively predict perceived authenticity of similarly-prejudiced statements | 1 – 3, 5 – 8 |
| 2 | Vicarious justification; justification by rearticulation | Self-reported prejudice should uniquely predict prejudice statements against the same target group; conversely, there should be no relationship between prejudice and the authenticity of other statements | 1 and 2 |
| 3 | Social projection | Prejudice should positively predict perceived descriptive normativity, which will in turn cause people to perceive prejudice statements as authentic | 3 and 4 |
| 4 | Prescriptive norms; vicarious justification | The positive relationship between prejudice and perceived authenticity should only be present when the prejudice is prescriptively non-normative | 3 – 6 |
| 5 | Balance | Prejudice should predict perceived authenticity more weakly when authenticity is portrayed negatively than when portrayed positively | 7 |
| 6 | Motivated reasoning | Participants motivated to be accurate in their perceptions of others should display a weaker relationship between prejudice and perceived authenticity than those motivated to express themselves | 8 |

Table 2

*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 3

*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 | N = 1,386 | .22 | [.18, .26] |

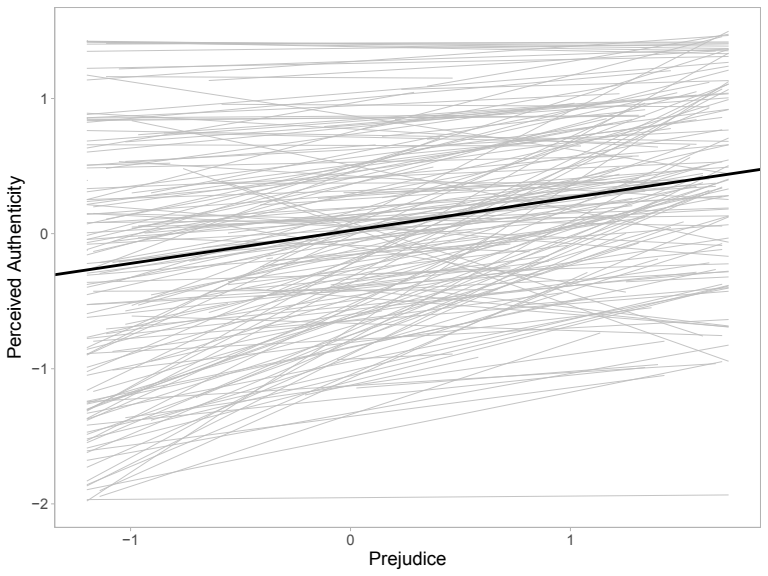
*Figure 1.* Prejudice against Muslims and politicians correlates positively with perceived authenticity of prejudiced statements toward the same groups (Study 1).



*Figure 2.* Negative statements about illegal immigrants are positively predicted by illegal immigrant prejudice; the same is true for Kansas State students.Unlike the regression equation that tested the hypothesis, these slopes are for presentational purposes and are *not* controlling for the influence of the other prejudice (Study 2).



*Figure 3.* Prejudice positively predicts perceived authenticity, across ten target groups. The thick, black line represents the average slope; the thin, grey lines represent slopes for each individual (Study 3).



*Figure 4.* The positive relationship between prejudice and perceived authenticity is stronger when the prejudice is perceived as prescriptively non-normative than when it is normative (Study 3).



*Figure 5.* There was no difference in perceived authenticity between conditions. The violin shapes represent the density of perceived authenticity. The solid dots and error bars represent the means and their 95% confidence intervals, respectively (Study 4).



*Figure 6.* Prejudice positively predicts authenticity in the suppression condition but not the expression condition (Studies 5 and 6).



*Figure 7.* The relationship between prejudice and perceived authenticity does not depend on the valence of how authenticity is portrayed (Study 7).



*Figure 8.* The relationship between prejudice and perceived authenticity does not depend on motivational goals (Study 8).

