Freedom of Racist Speech: Ego and Expressive Threats

Mark H. White II and Christian S. Crandall
University of Kansas

Do claims of "free speech" provide cover for prejudice? We investigate whether this defense of racist or hate speech serves as a justification for prejudice. In a series of 8 studies ($N = 1,624$), we found that explicit racial prejudice is a reliable predictor of the "free speech defense" of racist expression. Participants endorsed free speech values for singing racists songs or posting racist comments on social media; people high in prejudice endorsed free speech more than people low in prejudice (meta-analytic $r = .43$). This endorsement was not principled—high levels of prejudice did *not* predict endorsement of free speech values when identical speech was directed at coworkers or the police. Participants low in explicit racial prejudice actively avoided endorsing free speech values in racialized conditions compared to nonracial conditions, but participants high in racial prejudice increased their endorsement of free speech values in racialized conditions. Three experiments failed to find evidence that defense of racist speech by the highly prejudiced was based in self-relevant or self-protective motives. Two experiments found evidence that the free speech argument protected participants' own freedom to express their attitudes; the defense of other's racist speech seems motivated more by threats to autonomy than threats to self-regard. These studies serve as an elaboration of the Justification-Suppression Model (Crandall & Eshleman, 2003) of prejudice expression. The justification of racist speech by endorsing fundamental political values can serve to buffer racial and hate speech from normative disapproval.

Keywords: free speech, justification-suppression model, prejudice, reactance

Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.
First Amendment, U.S. Constitution, 1791.

Racists Getting Fired is a blog dedicated to getting people fired for making prejudiced remarks online. Racist posts on social media and the poster's employment information are put on the blog; visitors are encouraged to lobby the employer to fire these people over their statements. The blog includes a "Gotten" tab, where the moderators post examples of people they have successfully had fired (Haggerty, 2014).

Employment consequences for racist speech are widely spread through social media. An intern lost her position for posting a photo of her and a friend picking cotton with the caption, "Our inner [n-word] came out today" (Spata, 2015); when people started leaving a graduation ceremony early, a principal was fired for saying, "Look who's leaving! All the Black people" (Hanson, 2015); a firefighter was fired for saying that Dylan Roof, the man who shot and killed nine Black churchgoers in Charleston, "needs

to be praised for the good deed he has done" (Hensley, 2015); and a public school teacher was fired for tweeting statements like, "I am way too racist to be a teacher" (Bateman, 2015). A slew of police officers have also been fired or otherwise disciplined for expressing anti-Black prejudice on social media: likening Black protesters to "Planet of the Apes," (Workneh, 2015), or taunting a Black video game player by telling him, "I get paid to beat up [n-words] like you" (Plunkett, 2015; also see Frolik & Gokavi, 2015; Norman, 2015; Shipps, 2015).

Public response to these firings is mixed. Some claim that terminating an employee or expelling a student for prejudice violates the right to freedom of speech (e.g., Hongo, 2015; Pearce, 2015; Randazza, 2015; Reynolds, 2015; Volokh, 2015). The balance between promoting open speech and dissent is essential to a well-functioning civil democracy (Farber, 1991; Rosenfeld, 2002), and serious consequences for prejudicial speech may inhibit this good effect (Posner, 2002).

Students at predominately White institutions have recently challenged racist speech, feeding a "national conversation" about the tension between eradicating racism and ensuring freedom of speech. This conversation has turned into "speech wars" (Millhiser, 2015); some say activists who want a "safe space" from racism are coddled, closed-minded, and prefer conformity and comfort at the cost of freedom of speech (e.g., LoBianco & Scott, 2015; Lukianoff & Haidt, 2015; O'Neill, 2014; Rubin, 2015; Soave, 2015). Others say these appeals to free speech are coded ways to perpetuate racist oppression, and that freedom of speech does *not* imply freedom from the consequences of one's speech (Case & Weddington, 2015; Cobb, 2015; Gay, 2015; Goldberg, 2015; Manne & Stanley, 2015).

We consider what occurs when people observe someone else being punished for racist speech. Do prejudiced people strategi-

Correspondence concerning this article should be addressed to Mark H. White II, Department of Psychology, University of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66045. E-mail: markhwhiteii@ku.edu

cally use freedom of speech as a justification for—or defense against—these punishments for racism? We hypothesize that they do. What motivates prejudiced people to use free speech as a justification? We consider two hypotheses: (a) Learning someone else was punished for a prejudice that one shares threatens one's self-image, and (b) Seeing someone else punished threatens once sense of freedom, triggering reactance.

The Justification–Suppression Model, and Two Extensions

The conflict between the negative social value of prejudicial speech and the positive value of free expression creates a tension between suppression and expression. This tension is modeled by the justification-suppression model of the experience and expression of prejudice (JSM; Crandall & Eshleman, 2003). The JSM considers expressions of prejudice to be the result of three interdependent processes: genuine prejudice, suppression, and justification. Genuine prejudice is the unmanaged, unmanipulated internal representation of (negative) feelings about a group; these feelings have motivational force. Suppression processes (e.g., egalitarian values, social norms) inhibit the overt expression of prejudice, which helps avoid the experience of guilt, shame, or social punishments.

Suppressors diminish the *expression* of prejudice, but they do not reduce the underlying prejudicial affect. Prejudice—like any affect—has motivational force (Brehm, 1999). This creates a conflict between the desire and the disincentives to express; *justifications* reduce this conflict. A justification is a "psychological or social process that can serve as an opportunity to express genuine prejudice without suffering external or internal sanction" (Crandall & Eshleman, 2003, p. 425); justifications (e.g., attributions of responsibility, negative stereotypes, belief in a just world) allow otherwise inhibited prejudices to be expressed, free of guilt and social punishment. Suppressors keep people from both communicating their prejudices and consciously experiencing it to themselves, and justifications free both the expression of prejudice to others and the experience of it to themselves.

When someone is charged with inappropriate expression, justifications can restore one's social honor in many ways. Justifications can normalize discrimination ("but those people are dangerous!"); they can change the interpretation of a situation ("it was really not her fault"); or they can connect the apparently unacceptable behavior to a cherished value ("she's exercising free speech").

Vicarious Justification

The JSM was conceived as an intrapsychic model; it describes how people release expression of their own suppressed prejudices. But people provide explanations, make excuses, and justify the behavior of others as well (Skitka, 2009; Snyder, Higgins, & Stucky, 1983). This can be motivated by self-justification, sympathy and concern, or consistency and balance (Allen & Caillouet, 1994; Crandall, Silvia, N′Gbala, Tsang, & Dawson, 2007; Snyder & Higgins, 1988), among many reasons.

The social justification of other's prejudice was implicit in Crandall and Eshleman (2003), but here we make the notion explicit. We propose that people will present or endorse justifications for *another's* expression of prejudice as a direct function of

their own prejudice. When another person behaves in a discriminatory manner, or makes a socially contentious statement, people will justify that action or expression when it is called into question, provided that they share that prejudice and the situation calls for a justification.

Delegitimizing Suppressors

Justifications can release prejudice without sanction by *delegitimizing* a Suppressor.

A justification process can work by neutralizing the social or psychological power of a suppressor; when the social norm, value, or audience that works as a suppressor loses its potency, expression can occur unhindered and without cost. Instead of directly making a prejudice acceptable, it disables the suppression. The ultimate consequence, in either case, is the open expression of prejudice.

By concluding (or claiming) that the suppressive forces that silence prejudice are illegitimate, one sidesteps the content of the expression. The delegitimization of suppression by appeal to the safer ground of rights, liberties, and values provides a safe and principled foundation for messages or behaviors that can otherwise be difficult to justify. By pointing to free speech as a value, one neatly evades the racist speech itself.

In these studies, people are offered an opportunity to endorse a fundamental American value—free speech—to justify prejudice by delegitimizing the "politically correct" forces of suppression. This justification has the advantage of being both values-based and normative, and glosses over the fact that this argument makes anti-Black stereotypes and overt discrimination acceptable. We predict that people will justify the expression of another's prejudice by delegitimizing the source of suppression. This justification will occur as a direct function of their own prejudice.

Experiencing Vicarious Suppression: Two Hypotheses

What motivates people to justify another's racial prejudice? The internal affective conflict in the JSM provides the motivational power for justification. People have feelings that they wish to express, and desire to do so without guilt or social sanction. In extending to the justification of others, internal conflict can provide some impetus, but rarely enough to publicly express support for a contentious attitude.

We propose two kinds of alternative explanations. First, because the self is implicated when someone expresses a view congenial to one's own, one will defend another as an extension of the self; the justification is self-protective, maintaining one's self-regard. We call this the *ego threat hypothesis*. Second, because the punishment another receives for expressing an attitude threatens one's own freedom to express an attitude, the justification of others is motivated by an assertion of autonomy, by reactance due to a lost freedom (Brehm, 1966). We call this the *expressive threat hypothesis*. We compare these two hypotheses in the context of using the value of free speech as a justification for the expression of racial prejudice.

Ego Threat Hypothesis

The *ego threat hypothesis* is that vicarious suppression threatens the self-integrity of prejudiced people, which motivates them to

justify another's prejudice in an effort to restore their self-regard. People do not want to appear prejudiced (Devine, Monteith, Zuwerink, & Elliot, 1991; Dunton & Fazio, 1997; Dutton, 1971, 1976); it is normatively inappropriate to be explicitly prejudiced. People may harbor beliefs that, "The whole Black Lives Matter movement is misguided and out of hand" and believe that activists are simply "egotistical," without seeing themselves as prejudiced. But reading a news story about a college student being pilloried as racist on social media and fired from her job for posting this very sentiment online (Klausner, 2015) communicates that these beliefs are indeed prejudiced, non-normative, and subject to punishment. We hypothesize that this suppression makes those with similar beliefs feel prejudiced, which threatens their self-integrity.

People's sense of self-goodness depends upon meeting with others' approval. Leary and Baumeister (2000) propose that self-esteem represents a sense that one is valued and accepted by others. Self-affirmation theory (Sherman & Cohen, 2006; Steele, 1988) argues that self-integrity—seeing the self as "good and appropriate" (Sherman & Cohen, 2006, p. 186)—relies on knowing that one is living in accordance with cultural norms.

Feelings about one's self is linked to racial attitudes; ego threats increase negative racial attitudes, but affirming the self increases positive racial attitudes. Threats to the self can lead people to judge an ethnic minority job applicant more harshly, but affirming the self diminished discrimination (Fein & Spencer, 1997; Shapiro, Mistler, & Neuberg, 2010). Self-affirmed White participants report perceiving more racism in discriminatory actions (Adams, Tormala, & O'Brien, 2006) and more acceptance that White privilege exists in society (Knowles & Lowery, 2012; Knowles, Lowery, Chow, & Unzueta, 2014; Phillips & Lowery, 2015).

Seeing another lose a job for prejudicial remarks communicates to prejudiced people that their attitudes are unacceptable and subject to sanction. In our studies, the ego threat hypothesis predicts that the more anti-Black prejudice one harbors, the more one's sense of self is threatened by the job termination over anti-Black prejudice. To restore self-integrity, prejudiced people should defend their egos by justifying another's prejudice via free speech—the more the personal threat, the greater the justification through free speech values.

Expressive Threat Hypothesis

The *expressive threat hypothesis* is that vicarious suppression threatens the perceived expressive freedom of prejudiced people, and they can regain this sense of freedom by justifying another's prejudice. Vicarious punishment generates psychological reactance, which directly motivates attempts to restore this freedom— justification through free speech values reestablishes one's liberty and autonomy.

Reactance theory (Brehm & Brehm, 1981; Wicklund, 1974) posits that people believe they have free will, and if this freedom is threatened, they will act in a way to defend or regain that freedom. Reactance theory makes no claims as to why people value freedom, but it is a central Western cultural value (Wike & Simmons, 2015) and people do benefit when their behaviors are self-determined and consistent with what they believe (Ryan & Deci, 2000; Sheldon & Krieger, 2007).

This includes freedom of expression. People in individualistic cultures are motivated to and derive positive psychological bene-fits from expressing what they consider their "true" feelings and attitudes (Heppner et al., 2008; Kernis & Goldman, 2006; Kim & Sherman, 2007), especially attitudes that people cannot always express (Schlegel, Hicks, Arndt, & King, 2009). People suffer ill effects when their need for freedom is thwarted, and one response to this is to react against the agent that is frustrating one's need for freedom (Brehm & Brehm, 1981; Vansteenkiste & Ryan, 2013; Wicklund, 1974).

Can people feel reactance when they witness the censorship of another's speech? Wicklund & Brehm (1967; in Wicklund, 1974) had middle school students scheduled to hear a presentation on why the voting age should be lowered from 21 to 18 learn the day of that it had been cancelled either because the school board did not want the kids to hear the message or because the speaker was sick. Students in the former condition were more supportive of lowering the voting age to 18. The differential shift happened even when participants did not agree with the censored speech (Ashmore, Ramchandra, & Jones, 1971; reported in Brehm & Brehm, 1981). Censorship also increased one's desire to hear the censored message (Worchel & Arnold, 1973; Worchel, Arnold, & Baker, 1975).

People feel like their freedom is threatened when they witness someone else's freedom being threatened, a vicarious reactance (Andreoli, Worchel, & Folger, 1974; Sittenthaler, Traut-Mattausch, & Jonas, 2015; Sittenthaler, Jonas, & Traut-Mattausch, 2016). Students felt more threat to their own freedom when they witnessed another student being denied participation credit for an experiment because the researcher lost their packet (an illegitimate reason to deny someone their credit) than when denied credit for a legitimate reason (Sittenthaler et al., 2016). There is some reason to believe that vicarious reactance is a more cognitive response than self-reactance (which is primarily affective, Sittenthaler et al., 2016), but it is clear that people do feel reactance on the behalf of others.

Self-determination theory makes a stronger claim, that freedom (or autonomy) is a psychological need (Ryan & Deci, 2000). When people have freedom and are autonomously and self-determined to perform a behavior, they are internally motivated to do so. People can be internally or externally motivated to suppress prejudices (Plant & Devine, 1998); they can refrain from expressing prejudice because it is important to their self-concept (internal) or due to social pressures (external). Participants low in internal motivation (IMS) and high in external motivation (EMS) to suppress prejudice report more racial prejudice, feel pressure to comply with "politically correct" standards, and feel anger toward these standards. These "politically correct" norms elicit behavioral backlash among those high in EMS and low in IMS (Plant & Devine, 1998, 2001).

The JSM explicitly proposes that one consequence of prejudice suppression is psychological reactance (Crandall & Eshleman, 2003, p. 423). Prejudice reduction techniques that focus on social pressure to be unprejudiced sometimes have the ironic effect of increasing prejudice (Legault, Gutsell, & Inzlicht, 2011; see also Brauer, El-rafiy, Kawakami, & Phills, 2012). White Americans sometimes lament these "politically correct" norms, and resist them by depicting the White dominant group as victimized (Andreouli, Greenland, & Howarth, 2016; Augoustinos & Every, 2007; King, 2015). Those who believe that there are extremist "politically correct crusaders" in society also believe that free speech rights are eroding (Lalonde, Doan, & Patterson, 2000). The

expressive threat hypothesis leads us to argue that institutional punishments for prejudice threaten prejudiced people's expressive freedom, evoke reactance, and lead people to justify prejudice.

The Present Studies

We predict that prejudiced people will use freedom of speech as a justification for someone else's suppressed prejudice. We test this basic hypothesis in Study 1, conducted shortly after a highly publicized punishment for prejudiced speech. In Studies 2 and 3, we manipulate the context (an employee fired for anti-Black prejudice or a control condition), predicting that prejudiced people will marshal freedom of speech as a justification more when reading about an employee fired for anti-Black speech. We then examine why people might be motivated to justify another's punished prejudice. Studies 4–6 test the *ego threat* hypothesis, which argues that prejudiced people feel a threat to their sense that they are a good and moral person when they read about a similarly prejudiced person harshly punished for prejudiced speech. Studies 7 and 8 test the *expressive threat* hypothesis, which argues that prejudiced people feel a threat to their expressive freedom when they learn that someone has been punished for expressing a prejudice they also hold.

Study 1: Singing Racist Songs

In March of 2015, a video showing fraternity brothers in Sigma Alpha Epsilon (SAE) at the University of Oklahoma chanting a racist song on a bus went viral on the Internet. The school responded swiftly: The university president directly addressed the people in the video by saying, "You are disgraceful," the Oklahoma chapter of SAE was shut down, all fraternity members were forced to move out of their fraternity house within days, and the Oklahoma University Board of Directors expelled two students who were leading the chant on the bus (Chappell, 2015). We tested our most basic hypothesis—that the higher the level of prejudice, the greater the endorsement of free speech as a justification—days after this occurred. Higher levels of explicit racism should be associated with more agreement that the punishments violated the students' right to free speech.

Method

Participants. We recruited 176 participants with an Amazon Mechanical Turk[1] (MTurk) listing for a "Survey about a recent news event." Participants' ages averaged 38.49 years (*SD* = 13.13, ranging 19 to 77); the sample was 43.8% female, and 79.5% of participants identified as White/Caucasian. Participants were paid $0.50 to complete the study.

Procedure. The survey took place 6 days after the Oklahoma video went viral; 69.3% of participants had heard of the incident. Participants were given a "description of the recent event" after agreeing to participate in the study. All participants read a brief description about the SAE controversy and subsequent punishment from the University of Oklahoma. The racist song was sung to the tune of, "If You're Happy and You Know it, Clap Your Hands." The lyrics were that there will "never be a [n-word] in S-A-E," and also made reference to lynching with, "You can hang him by a

tree, but he'll never sign with me." Participants then completed a brief questionnaire.

Measures.

Free speech. Three items measured perceived violations of free speech on a seven-point scale (α = .94): "Kicking the fraternity off campus is a violation of their free speech," "The students in the video have a right to free speech, so they should not have been expelled," and "The university being so harsh on the students in the video is not respecting the students' freedom of speech."

Anti-Black prejudice.

Symbolic racism. Anti-Black beliefs were assessed with the symbolic racism scale (eight items; α = .92; Henry & Sears, 2002). An example item reads: "It's really a matter of some people not trying hard enough; if Blacks would only try harder they could be just as well off as Whites."

Specific prejudicial emotions. Participants were asked: "In general, how much do you feel the following emotions when you are around Black people?" (see Dijker, 1987). A list of 12 emotions followed, and participants were asked to respond on a seven-point scale anchored from 1 (*not at all*) to 7 (*very much so*). A maximum likelihood, direct oblimin rotation exploratory factor analysis of these items yielded two factors: positive and negative emotions. The positive scale (α = .94) included happiness, admiration, affection, respect, warmth, and relaxed. The negative scale (α = .93) included: annoyed, anger, fear, distrust, anxiety, and discomfort.

Feeling thermometer. Participants were asked to indicate their "overall feeling about the following groups based on a scale from 0 (*very cold and negative feelings*) to 100 (*very warm and positive feelings*)." Ten groups were listed in a random order, including alcoholics, fat people, politicians, and elderly people. Their response to Black people was used to measure anti-Black prejudice.

Overall prejudice. We standardized each of the four prejudice measures, reverse-scored the positive emotions scale and feeling thermometer, and averaged them together to create an overall index of anti-Black prejudice (α = .76).

Ideological variables. We also assessed social dominance orientation (SDO; Sidanius & Pratto, 1999) with six items (α = .89). Political identification was measured by averaging two Likert scales anchored with 1 (*liberal*) to 7 (*conservative*) and 1 (*democrat*) to 7 (*republican*; *r* = .88).

Results

Anti-Black prejudice positively correlated with freedom of speech, *r* = .47, *p* < .001. Each of the constituent measures correlated with freedom of speech: symbolic racism, *r* = .46; feeling thermometer, *r* = −.37; positive affect toward Black people, *r* = −.35; negative affect toward Black people, *r* = .25, all *p*s ≤ .001.

SDO, political identification, and gender (0 = female, 1 = male) were also significantly related to the free speech items, *r*s = .44, .24, and .26, respectively. To see whether any of these accounted for the relationship between prejudice and free speech relevance, we entered these three predictors and overall prejudice

[1] We used exclude functions on TurkPrime.com (Litman, Robinson, & Atterbock, 2016) to ensure that no one participated in more than one study in this paper.

simultaneously into a linear regression model. The overall model was significant, $R^2 = .28$, $F(4, 170) = 16.66$, $p < .001$. All predictors were significant, including overall prejudice, $b = .54$, $SE = .18$, $t(170) = 3.08$, $p = .002$; SDO, $b = .43$, $SE = .14$, $t(170) = 3.00$, $p = .003$; and gender, $b = .71$, $SE = .28$, $t(170) = 2.53$, $p = .012$, excepting political identification, $b = -.01$, $SE = .10$, $t(170) = -0.09$, $p = .926$.

Discussion

The more prejudice participants reported, the more they agreed that expelling the students and kicking the fraternity off campus violates free speech. This relationship held regardless of how prejudice was measured, and it held after statistically controlling for gender, political ideology, and SDO.

These data are consistent with the hypothesis that "freedom of speech" can be used as a justification for prejudice. We do not know if an as-yet-unmeasured third variable creates the correlation between prejudice and support for freedom of speech; we address this issue experimentally in Studies 2 and 3.

Studies 2 and 3: Justification or Principle?

If a principled belief in the sanctity of freedom of speech is naturally related to racial prejudice, then prejudiced people should endorse the relevance of freedom of speech equally *regardless* of context. If the deployment of freedom of speech defense is *not* principled but rather a justification, it should be especially appealing in contexts where it serves as a prejudice justification. People low in prejudice will similarly recognize the justification value of free speech, but they should increasingly avoid endorsing it. In Studies 2 and 3, we manipulate whether or not the context is conducive to free speech being a justification for prejudice.

In Study 2, we experimentally manipulated the offending speech to be anti-Black or anti-police. Anti-Black prejudice should correlate with free speech only when the punished speech is anti-Black, and not when the punished speech is anti-police. In Study 3, we manipulated a news story to be about an employee fired for anti-Black prejudice or for a completely unrelated story.

Study 2 Method

Participants. We recruited 251 participants from MTurk to complete the study. Participants' ages averaged 34.88 years ($SD = 11.43$, ranging 18 to 73); the sample was 47.4% female, and 78.1% of the participants identified as White/Caucasian. Participants were paid $0.60 to complete the study. To determine the sample size, we simulated a dataset where one condition had a correlation between free speech and prejudice at $r = .47$ (the effect size found in Study 1) and the other at $r = .00$. We used the effect size of the interaction as an effect size estimate for the current study and collected enough data to ensure 95% power.

Procedure. Participants read a "description of a recent news event" after agreeing to participate in the study. Participants were randomly assigned to one of two conditions: an *anti-Black* or an *anti-police* condition. The passages read:

> Colin Slator, the director of content and advertising at a prominent phone company, was fired after management discovered comments he made about [Black people/police officers] on Facebook.

> In recent months, there have been many demonstrations in the United States protesting the use of violence and deadly force by the police against African Americans. There was a protest like this in the town Colin lives in. He posted on Facebook how he felt about the issue:

> "These [protesters/cops] are just a bunch of [looters and thugs/racists and pigs]. They're all bastards. [Blacks/police] are the ones causing all of this racial tension in America right now, and I'm sick of it. Fuck them."

Someone saw this post and forwarded the comments to management at Slator's job. They decided to fire Slator, saying that he did not "represent the values that our company stands for."

Participants were instructed to read this passage carefully and then complete a questionnaire.

Measures.

Free speech. The relevance of freedom of speech to the firing was assessed with three items: "Management's actions went against Colin Slator's freedom of expression," "Colin Slator's bosses disrespected his right to free speech," and "Firing Colin Slator is a violation of his rights to free speech" ($\alpha = .95$).

Anti-Black prejudice. We measured prejudice using Henry and Sears's (2002) symbolic racism scale and a feeling thermometer. We reverse-scored the thermometer, standardized both measures, and averaged them together, $r = .45$, $p < .001$[2] to measure anti-Black prejudice. We use slightly different measures of prejudice across the studies to increase the robustness of measuring the construct (Crandall & Sherman, 2016, pp. 95–96), and we employed fewer measures of prejudice in subsequent studies to reduce costs.

Study 2 Results

We hypothesized that anti-Black prejudice would only predict the relevance of freedom of speech for the firing in the condition where the ex-employee was fired for anti-Black remarks.

We tested this hypothesis by entering prejudice (standardized), condition (mean-centered), and the interaction between the two in a linear regression model predicting freedom of speech. The overall model was significant, $R^2 = .13$, $F(3, 247) = 12.79$, $p < .001$. There was a main effect of prejudice, $b = .42$, $SE = .11$, $t(247) = 3.88$, $p < .001$ and condition, $b = -.49$, $SE = .22$, $t(247) = -2.26$, $p = .025$.

The interaction that tested the primary hypothesis was also significant, $\Delta R^2 = .08$, $b = 1.01$, $SE = .22$, $t(247) = 4.64$, $p < .001$. We examined the simple slopes predicting free speech with prejudice at both levels of the independent variable (see Figure 1). Prejudice predicted *greater* freedom of speech relevance in the *anti-Black* (coded 1) condition, $b = .93$, $SE = .16$, $t(247) = 5.74$, $p < .001$. In the *anti-police* (coded 0) condition, prejudice did not predict free speech, $b = -.09$, $SE = .15$, $t(247) = -0.58$, $p = .563$. Probing this interaction further using the Johnson-Neyman technique (Hayes, 2013), we found that participants scoring $Z = 1.12$ and higher in prejudice thought free speech was violated more in the *anti-Black* condition than the *anti-police* condition; participants scoring $Z = 0.06$ and lower in prejudice thought free speech

---

[2] Results were identical in direction and statistical significance when doing analyses with each of the measures of prejudice separately.
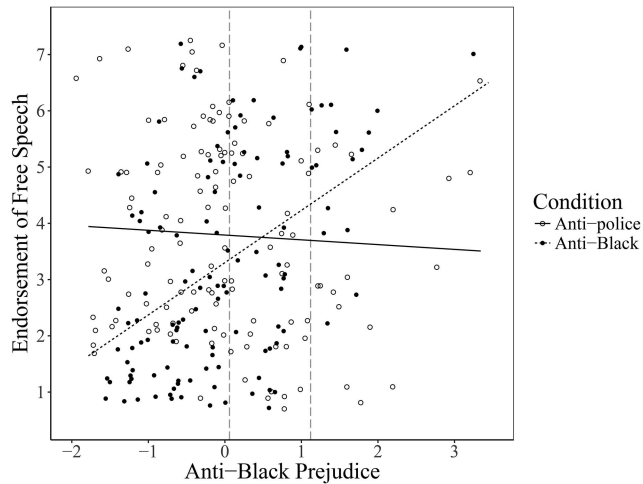
Figure 1. Anti-Black prejudice (standardized) predicted perceived violation of free speech, but only in the anti-Black condition (Study 2). Participants who fall on the outside of the vertical lines differ by condition in their endorsement of freedom of speech.

was violated more in the *anti-police* condition than in the *anti-Black* condition.

The anti-police comparison condition may attenuate a preexisting relationship between prejudice and free speech; we addressed this limitation in Study 3 by exposing participants in a control condition to an article with no mention of job termination or offensive speech.

Study 3 Method

Participants. We recruited 245 participants from MTurk to complete the study. Participants' ages averaged 34.35 years (SD = 10.56, ranging 18 to 71); the sample was 49.8% female, and 80.8% of the participants identified as White/Caucasian. Participants were paid $0.75 to complete the study. Sample size was determined using the same power analysis as in Study 2.

Procedure and materials. We reduced the level of hate to expressions of intergroup anxiety, which tests the generalizability of operationalization and whether a similar effect will emerge for less intense speech. Participants were asked to read a "news story" after agreeing to participate in the study. Participants were randomly assigned to one of two conditions: a *fired* condition or a *control* condition. The *fired* condition told of how an employee was fired for writing:

> Black people "are so touchy about race that it is difficult to get along with them." He said that Black people could "be combative" and "assume the worst from White people." He concluded by saying that this "makes me feel uncomfortable sometimes, which is why I don't really like to be around them.

The *control* condition was a story about how the French macaron pastry is supplanting the cupcake as America's new chic dessert.

Measures.

Free speech values. Perceptions of how much firing an employee for a post on the Internet was assessed with three items:

"Free speech protects people from being fired for something they said on the Internet," "Freedom of expression means people should not be punished for something they say on the Internet," "Free speech means that people should be able to express themselves without facing severe consequences for their opinion" (α = .84).

Anti-Black prejudice. We assessed prejudice using Henry and Sears's (2002) symbolic racism scale and a set of five feeling thermometers: Black people, "Black Lives Matter" activists, civil rights leaders, the NAACP, and the Black Panther Party. We standardized both measures, reverse-scored the mean of the five thermometers, and averaged the two scales together as our measure of prejudice, r = .70, p < .001.[3]

Study 3 Results

We hypothesized that anti-Black prejudice would predict the violation of freedom of speech values for firing someone for an online post stronger in the *firing* condition than the *control* condition. We tested this by entering prejudice (standardized), condition (mean-centered), and their interaction in a linear regression model predicting freedom of speech. The overall model was significant, R^2 = .12, $F(3, 241)$ = 10.67, p < .001. There was a main effect of prejudice, b = .46, SE = .10, $t(241)$ = 4.71, p < .001, but not condition, b = −.21, $t(241)$ = −1.09, p = .276.

The interaction that tested the primary hypothesis was significant, ΔR^2 = .03, b = .53, SE = .19, $t(241)$ = 2.71, p = .007. Prejudice was a significant predictor of free speech in the *firing* (coded 1) condition, b = .72, SE = .14, $t(241)$ = 5.16, p < .001, but not in the *control* (coded 0) condition, b = .19, SE = .14, $t(241)$ = 1.39, p = .165. Probing this interaction further with the Johnson-Neyman technique (Hayes, 2013), participants scoring Z = 2.09 and higher in prejudice thought free speech is violated more in the *firing* condition than the *control* condition; participants scoring Z = −0.37 and lower in prejudice thought free speech is violated more in the *control* condition than in the *firing* condition (see Figure 2).

Studies 2 and 3 Discussion

In both studies we found that objections based on free speech were unprincipled—people high in racial prejudice brought their values out only when the speech was racialized. By contrast, people low in prejudice backed away from endorsing free speech values in these racialized contexts—especially in Study 3, where the effect of condition on free speech endorsement was more prevalent for people low in prejudice than those high (we return to this in the General Discussion). This could be a case of undersampling highly prejudiced participants, given that demographic groups who express fewer prejudices are overrepresented in MTurk samples (Huff & Tingley, 2015). If perceived violations of freedom of speech were a consistent moral principle that is higher among those with racial prejudice, then it should be equally endorsed regardless of context. Instead, free speech concerns were endorsed only when they justified prejudice.

Free speech values emerge consistently as justifications. But in the JSM, justifications are conceived as releasers for *one's own*

---

[3] Results were virtually identical in direction and statistical significance when doing analyses with each of the measures of prejudice separately.
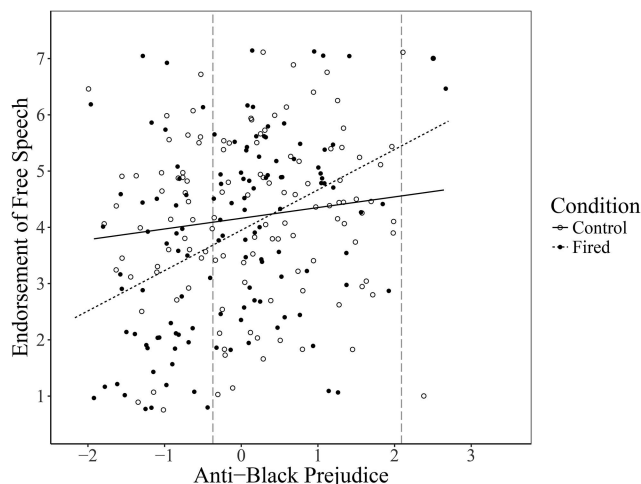
Figure 2. Anti-Black prejudice (standardized) predicted perceived violation of free speech, but only in the fired condition (Study 3). Participants who fall on the outside of the vertical lines differ by condition in their endorsement of freedom of speech.

prejudice (Crandall & Eshleman, 2003); why do people justify the prejudiced speech of someone else? This constitutes the task for the remainder of this paper. In the next five studies we consider the two main hypotheses: (a) is suppression motivated by a threat to ego (Studies 4 – 6) or (b) is suppression motivated by threats to expression (Studies 7 and 8)?

Studies 4–6: The Ego Threat Hypothesis

People generally want to feel good about themselves (or people from Western cultures do; Crocker & Wolfe, 2001; Heine, Lehman, Markus, & Kitayama, 1999; Leary & Baumeister, 2000; Sherman & Cohen, 2006; Steele, 1988; Tesser, 2000). The ego threat hypothesis is that prejudiced people feel a threat to their self-regard when they hear about someone getting fired for expressing their attitudes; it indicates their attitudes will be punished. To affirm their self-integrity (Sherman & Cohen, 2006; Steele, 1988), prejudiced people call upon normative values—like freedom of speech—to justify the prejudiced speech of people like them. We test this hypothesis across three studies, employing three different experimental manipulations of self-regard and self-relevance.

Participants in Studies 4–6 read about someone getting fired for prejudiced speech. In Study 4, we instructed participants to complete either an affirming- or control-task. We predicted that affirming the self would attenuate the relationship between prejudice and freedom of speech, because affirming the self mitigates the threat brought about by reading about harsh suppression of prejudice. In Study 5, the ego threat was in the domain of prejudice; we threatened participants by making them feel prejudiced. We predicted that people would endorse the free speech justification more when they are primed to feel prejudiced. In Study 6, we manipulated how similar participants felt toward the employee who was punished for prejudiced speech, and we predicted that this increased self-relevance of the punishment would cause greater endorsement of the free speech justification.

We manipulated self-affirmation or self-threat in Studies 4 and 5, but did not measure it as a mediating variable, because merely filling out these types of questions can be self-affirming (Kimble, Kimble, & Croy, 1998; Steele, Spencer, & Lynch, 1993), which could obviate the need for participants to use freedom of speech as a justification.

Study 4 Method

Participants. We recruited 206 participants[4] form MTurk to complete the study. Participants' ages averaged 34.45 years ($SD =$ 12.16, ranging 18 to 83); the sample was 49% female, and 80.1% of participants identified as White/Caucasian. Participants were paid $0.70 to complete the study.

Procedure. All participants first read about someone who was fired for saying disparaging remarks about Black people on social media. The study was ostensibly about memory, and participants were instructed to read the passage carefully because they would be answering questions about it later in the study. The passage was similar to the passage in Studies 2 and 3, but with different prejudiced speech:

> He posted a link to a story covering a racial protest in his city and wrote, "I'm just going to go ahead and say it . . . the Blacks are the ones causing the problems and this 'racial tension.' I guess that's what happens when you flunk out of school and have no education." He went on to say later in the post that, "This kind of trouble is why people don't want to drive through the Black part of town."

The story concluded with the employee being fired for this post.

Participants were then randomly assigned to complete a self-affirmation manipulation or a control task. In the *self-affirmation* condition, we employed a widely used self-affirmation manipulation by asking participants to select a "value, quality, or aspect" of life that is "most important to you in your life" from a list (e.g., sense of humor, creativity, romance, relations with friends and family; see McQueen & Klein, 2006). Participants were then prompted to write about why it is so important to them and to write about a time the value they selected made them feel good about themselves.

In the *control* condition, we asked participants to select a common household item from a list (e.g., pencil, shoelaces, ironing board; see Harvey & Oswald, 2000). Participants were then prompted to write down five uses for this household item and to describe where it would be located in a superstore.

Participants then answered three free speech items: "Firing Colin Slator is a violation of his rights to free speech," "Colin Slator has a right to free speech, so he should *not* have been fired for his Facebook post," "The company being so hard on Colin Slator is not respecting his freedom of speech" ($\alpha = .93$) and the same four prejudice scales ($\alpha = .81$) as in Study 1.

Study 4 Results and Discussion

We hypothesized that bolstering participants' self-integrity would attenuate the relationship between prejudice and free

---

[4] Because of funding availability, we collected fewer data for this study, but still enough for 90% power based on the power analysis described in Study 1.

speech. We tested this hypothesis by entering in condition (mean-centered), prejudice (standardized), and the prejudice by condition interaction simultaneously in a linear regression equation.

The overall model was significant, $R^2 = .13$, $F(3, 202) = 9.74$, $p < .001$. There was a main effect of prejudice, $b = .66$, $SE = .13$, $t(202) = 5.04$, $p = .012$, but no main effect of condition, $b = .13$, $SE = .26$, $t(202) = 0.50$, $p = .617$. The interaction was *not* significant, $\Delta R^2 = .01$, $b = .39$, $SE = .26$, $t(202) = 1.49$, $p = .139$. In addition to this lack of significance, the pattern of coefficients was opposite than our prediction: Prejudice significantly predicted freedom of speech in both the *control* condition (coded 0), $b = .47$, $SE = .19$, $t(202) = 2.52$, $p = .012$, and *self-affirmation* condition (coded 1), $b = .86$, $SE = .18$, $t(202) = 4.78$, $p < .001$, with a slightly higher coefficient in the self-affirmation condition, where we predicted attenuation.

The *ego threat hypothesis* did not gain support in Study 4; protecting participants from threat did not affect the relationship between prejudice and the freedom of speech justification. However, the self-affirmation manipulation was unrelated to prejudice; nobody in the self-affirmation condition wrote about their non-prejudiced attitudes to affirm themselves.

The effect may only emerge within the domain of prejudice, making a global self-affirmation manipulation ineffective at buffering the ego threat. We used a manipulation in Study 5 to address this possibility, creating a specific threat within the domain of prejudice. We hypothesized that priming participants to feel prejudiced would increase their use of the free speech justification for prejudice.

Study 5 Method

Participants. We recruited 134 participants from MTurk to complete the study. Participants' ages averaged 32.16 years ($SD = 10.24$, ranging 18 to 68); the sample was 40.3% female, and 79.9% of participants identified as White/Caucasian. Participants were paid $0.75 to complete the study. In Studies 5, 6, and 7, we recruited enough participants to ensure 80% power when Cohen's $d = .5$ for the hypothesized main effect of condition on endorsement of free speech.

Procedure. Participants were randomly assigned to complete one of two scales. Each condition presented a list of 14 social groups that Americans generally see as unacceptable to express prejudice against (e.g., Black people, Asians, immigrants, gays and lesbians; Crandall, Eshleman, & O'Brien, 2002; Crandall, Ferguson, & Bahns, 2013). In a *threat* to nonprejudiced self-image condition, participants were asked to indicate if they have *ever* said anything negative about each group, marking "yes" or "no" for each. In a *control* condition, participants were asked to indicate if they *frequently* say negative things about this group (see Monin & Miller, 2001; Salancik & Conway, 1975). Participants were asked to tally up how many times they indicated "yes" and write it in a space below the scale. Only 6% of participants in the *threat* condition indicated "Yes" zero times ($M = 2.04$); 47% did in the control condition ($M = 7.65$), $\chi^2(1) = 28.65$, $p < .001$. In an unrelated study from our lab, this manipulation significantly increased negative feelings about the self (e.g., disappointed with myself; Devine et al., 1991) for participants in the *threat* condition, $t(195) = 3.51$, $p = .001$, Cohen's $d = .44$ [.15, .72].[5]

Participants read the same story as in the fired condition in Study 3. Right before completing the questionnaire, participants were again asked to write how many times they marked "yes" in the aforementioned scale, as to prime them with the threat (or lack thereof) again. We employed the same freedom of speech items as in Study 4 ($\alpha = .92$). We measured prejudice by averaging together three scales ($\alpha = .84$): symbolic racism (Henry & Sears, 2002), social distance (Crandall, 1991), and overall negative affect felt toward Black people (2 items: 1, *not at all positively* to 5 *extremely positively*, reverse-scored; and 1, *not at all negatively* to 5 *extremely negatively*).

Study 5 Results and Discussion

We hypothesized that threatening participants' nonprejudiced self-image would cause them to justify someone else's punished prejudice using freedom of speech. The *threat* condition ($M = 3.99$, $SD = 1.96$) and *control* condition ($M = 4.19$, $SD = 1.85$) did not differ from one another in the perceived violation of free speech, $t(132) = -0.63$, $p = .532$, $d = -.11$ [$-.44$, .23]. Prejudice and condition had no interactive effect on free speech, $b = -.03$, $SE = .73$, $t(130) = -0.05$, $p = .958$, but prejudice again correlated with perceived violation of free speech, $r = .43$, $p < .001$.

The data do not support the *ego threat hypothesis*; threatening participants' selves within the domain of prejudice did not affect how much they chose to justify someone else's suppressed prejudice. Participants' who were in a control condition justified the prejudiced speech just as much as those who were induced to feel prejudiced.

In a third attempt to test for ego threat effects, in Study 6 we manipulated how similar participants felt to the terminated employee. Feeling similar to another can increase the inclusion of them in one's self-concept (Aron, Aron, Tudor, & Nelson, 1991; Cialdini, Brown, Lewis, Luce, & Neuberg, 1997; Smith, Coats, & Walling, 1999), and people are more likely to feel how others feel when they are similar to them (e.g., Batson, Turk, Shaw, & Klein, 1995; Krebs, 1975; Preston & de Waal, 2002). We thus hypothesized that similarity would increase justification.

Study 6 Method

Participants. We recruited 135 participants from MTurk to complete the study. Participants' ages averaged 32.03 years ($SD = 9.90$, ranging 19 to 64); 54.1% of the sample was female, and 77% of participants identified as White/Caucasian. Participants were paid $0.75 to complete the study.

Procedure. Participants were asked to read a description of a man named Colin Slator carefully and were randomly assigned to either a *similar* or *different* condition. The passage contained simple information about Colin: where he is from, where he went to college, what type of music he likes, what his friends say about him, what pets he has, what food he likes, and so forth. In the *similar* condition, participants were asked to think about all the ways they were similar to Colin while reading the passage, and we prompted them to list five ways that they were similar to him

---

[5] All confidence intervals reported are 95% confidence intervals.

below the passage. In the *different* condition, participants were asked to think about how different they were from Colin, and they listed five ways that they were different from Colin below the passage.

On the next screen, participants completed a manipulation check: "How similar do you feel to Colin Slator?" and "How different do you feel that you and Colin Slator are from one another?", $r = -.70$, $p < .001$. Participants responded to these questions on five-point Likert scales, and we scored the items such that higher scores indicated feeling more similar to Colin. Participants in the *similar* condition ($M = 3.71$, $SD = 0.84$) reported more similarity to Colin than those in the *different* condition ($M = 3.35$, $SD = 0.87$), $t(133) = 2.46$, $p = .015$, $d = .42$ [.08, .76].

Participants then read about how Colin Slator was fired from his job for posting disparaging remarks about Black people on his personal Facebook page (the same remarks as in Studies 3 and 5). Participants then answered the same free speech items ($\alpha = .94$) and prejudice items ($\alpha = .79$) as in Study 5.

Study 6 Results and Discussion

We hypothesized that if participants felt similar to the ex-employee, they would justify his prejudice. However, participants in the *similar* ($M = 3.58$, $SD = 1.77$) and *different* ($M = 4.00$, $SD = 1.90$) conditions did not significantly differ on how much they endorsed the freedom of speech items, $t(133) = -1.31$, $p = .19$, $d = -.23$ [$-.56$, .11]. Moreover, the manipulation check (how similar participants reported feeling to Colin Slator) did not correlate with endorsement of the freedom of speech items, $r = -.07$, $p = .40$. There was no interaction between condition and prejudice on free speech, $b = .42$, $SE = .37$, $t(131) = 1.13$, $p = .259$, but prejudice again predicted freedom of speech, $r = .30$, $p = .001$.

Our third test of the *ego threat hypothesis* again failed to yield support for the relevance of self-threat to the present phenomenon: Participants who felt more similar to the terminated employee did *not* endorse the freedom of speech justification than those who felt dissimilar to him.

Ego Threat Discussion

The logic of the ego threat process is this: Prejudiced people justify someone else's prejudice because the other person's punishment makes them feel diminished—an attitude they share is punished as socially unacceptable, and this threatens their self-regard (Devine et al., 1991; Leary & Baumeister, 2000; Sherman & Cohen, 2006). To restore the self, prejudiced people endorse freedom of speech, which reframes the issue and justifies the behavior (Kristiansen & Zanna, 1994).

We found no support for this hypothesis across three studies; affirming general self-integrity, threatening self-integrity in the domain of prejudice, and increasing the similarity to the targeted employee all failed to have the predicted effects. Self-regard is integral to many interracial attitude processes (e.g., Adams et al., 2006; Fein & Spencer, 1997; Knowles et al., 2014), but we found no evidence that self-involvement, self-protection, or self-implication played an important role in justifying racist conduct through free speech values.

Each of these studies has its limits: Making people feel prejudiced might work only on those low in prejudice; news of racist

speech *after* a similarity manipulation may overwhelm a manipulation based on personality or food preferences. But taken together—all with null results, all trending in the opposite direction—change the balance of likelihood. One can hardly prove the null with three studies, but they do reduce the likelihood of the ego threat hypothesis, and encourage us to consider alternatives.

We focus on a different threat in our last two studies: The *expressive threat hypothesis*. Instead of these fired-for-prejudice stories threatening prejudice people's positive attitudes toward themselves, these stories might threaten prejudiced people's belief that they live in an environment that allows them to express their beliefs and attitudes.

Studies 7 and 8: The Expressive Threat Hypothesis

People believe that their behaviors are freely determined, and if this belief is threatened, people will try to recover this feeling of freedom (Brehm & Brehm, 1981). Expressing oneself is a behavior that people in individualistic cultures are strongly motivated to feel is self-determined, and they feel good when they express their "true" selves (e.g., Heppner et al., 2008; Kernis & Goldman, 2006; Kim & Sherman, 2007; Schlegel et al., 2009).

People experience psychological reactance in response to censorship (e.g., Wicklund, 1974), when others' freedoms are threatened (e.g., Sittenthaler et al., 2016), and external motivators to decrease prejudice can actually increase prejudice (e.g., Legault et al., 2011). We argue that harsh punishments of prejudiced speech—such as job termination—threatens the expressive freedom that similarly prejudiced people so desperately want to feel. One way to mollify this expressive threat might be to justify the prejudiced speech.

The *expressive threat* hypothesis states that punishments for another's prejudice threatens the expressive autonomy of similarly prejudiced people, which motivates their marshaling of a value (freedom of speech) to oppose the punishment and justify the expressed prejudice. We test this hypothesis in Studies 7 and 8.

In Study 7, we test whether priming autonomy threat alone will increase endorsement of free speech. We then examine in Study 8 how prejudice fits into this process; we test whether the relationship between anti-Black prejudice and perceived violation of freedom of speech is mediated by expressive autonomy threat, but only when the offensive speech is anti-Black. The pair of these studies constitutes an experimental causal-chain design (Spencer, Zanna, & Fong, 2005), which involves pairing a correlational relationship between the expressive threat and free speech in Study 8 with Study 7, where we manipulate expressive threat.

Study 7 Method

Participants. We recruited 128 participants from MTurk to complete the study. Participants' ages averaged 36.16 years ($SD = 12.80$, ranging 18 to 77); the sample was 51.6% female, and 75.8% of participants identified as White/Caucasian. Participants were paid $0.50 to complete the study.

Procedure. In a study ostensibly about "writing style and social attitudes," participants were randomly assigned to one of two conditions: an *expressive threat* or *control* writing prompt. The *expressive threat* prompt read:

Many people sometimes feel like they are not allowed to be themselves. There are some times that people feel like they cannot express their thoughts or ideas. Take a few moments to think about a time in your life when you felt like this—when you felt pressured to think what others thought or to have certain beliefs, or when you couldn't voice your opinion.

They were instructed to write about one of these experiences in detail in a text box below. The *control* condition prompted participants to write about their daily routine (Baldwin, Biernat, & Landau, 2015, Study 5). Participants then answered a 14-item "social attitudes" scale, where the three free speech values items from Study 3 were presented randomly throughout the scale ($\alpha = .65$).

Study 7 Results and Discussion

Writing about *expressive threat* condition ($M = 4.54$, $SD = 1.14$) increased participants' endorsement of free speech values compared to the *control* condition ($M = 4.01$, $SD = 1.56$), $t(126) = 2.22$, $p = .028$, $d = 0.39$ [.04, .74]. After remembering a time participants were restricted in their speech, they increased their agreement that free speech values protect people from harsh punishments for their speech. This causal relationship makes plausible the argument that feelings of expressive threat could motivate prejudiced people to justify another's prejudice. We test this hypothesis in Study 8. Participants were assigned to read a new story about a firing for *anti-Black* or *anticoworker* speech; we tested whether concerns about threats to autonomy mediate the relationship between prejudice and free speech values in both conditions. We predict that perceived expressive threat will be a strong mediator in the race-relevant anti-Black condition, but not in the anticoworker condition.

Study 8 Method

Participants. We recruited 348 participants from MTurk. Participants' ages averaged 34.94 years ($SD = 11.30$, ranging 19 to 71); the sample was 48.0% female, and 79.6% of participants identified as White/Caucasian. Participants were paid $0.75 to complete the study. We determined our sample size based on recommendations for moderated mediation models (Preacher, Rucker, & Hayes, 2007).

Procedure. Participants read a "news story" after agreeing to participate in the study. They were randomly assigned to read that an employee—named Connor Bell—was fired for posting *anti-Black* or *anticoworker* sentiments on his personal Facebook page. The *anti-Black* statements read:

Bell was at a coffee shop, and Black customers were also in the café. He posted on Facebook that these Black customers were being "so annoying." He wrote that Black people are "rude," "unintelligent," "too loud," and "lazy." Bell also posted that he is "just so tired of being around Black people."

In the *anticoworker* condition, Connor Bell lobbed the same insults at his bosses and coworkers. Participants then completed a questionnaire.

We employed the same free speech items as in Studies 4 through 6 ($\alpha = .97$), and used the symbolic racism scale ($\alpha = .89$; Henry & Sears, 2002) to assess prejudice. Our measure of *expressive*

threat was slightly adapted from the autonomy subscale of the basic psychological needs scale (Johnston & Finney, 2010): "I am free to decide for myself how to live my life," "I feel pressured to think what others want me to think," "I feel free to express my ideas and opinions," "I feel like I can pretty much be myself," "I feel free to be who I am," "I can voice my opinion," and "I feel controlled and pressured to have certain beliefs" ($\alpha = .92$). Items were scored such that higher scores indicated more threat to expression.

Study 8 Results and Discussion

We hypothesized that anti-Black prejudice would predict more expressive threat and more perceived violation of free speech, but only in the *anti-Black* condition (coded 1) and not in the *anticoworker* condition (coded 0). We regressed expressive threat and free speech on prejudice (standardized), condition (mean-centered), and their interaction in two separate linear regression models to test the corresponding hypotheses; coefficients are displayed in Table 1.

The prejudice by condition interactions were the critical tests of the hypotheses, and both were significant, $ps \leq .020$. We probed both of these interactions by examining the simple slopes of the dependent variables regressed on prejudice in each of the two conditions.

Expressive threat. Prejudice predicted more expressive threat in the *anti-Black* condition, $b = .36$, $SE = .09$, $t(344) = 3.90$, $p < .001$, but not in the *anticoworker* condition, $b = .06$, $SE = .09$, $t(344) = 0.71$, $p = .479$. This effect of condition on free speech was significant only for those high in prejudice; using the Johnson-Neyman technique (Hayes, 2013), we found that participants scoring $Z = 0.78$ and higher in prejudice felt more expressive threat in the *anti-Black* condition than the *anticoworker* condition. There were no differences between conditions for participants below this threshold.

Free speech. Prejudice predicted more violation of Free speech values in the *anti-Black* condition, $b = .97$, $SE = .13$, $t(344) = 7.62$, $p < .001$, and to a lesser extent in the *anticoworker* condition, $b = .31$, $SE = .12$, $t(344) = 2.52$, $p = .012$. The interaction is displayed in Figure 3, $\Delta R^2 = .03$, $b = .67$, $SE = .18$, $t(344) = 3.77$, $p < .001$. Using the Johnson-Neyman technique (Hayes, 2013), we found that participants scoring $Z = 0.55$ and higher in prejudice thought free speech was violated more in the

Table 1

Unstandardized Coefficients for the Regression Models Tested in Study 8

Outcome	Predictor	b	SE	t	p
Expressive threat	Condition	.09	.13	.67	.501
	Prejudice	.21	.06	3.34	.001
	Prejudice × Condition	.30	.13	2.33	.020
Freedom of speech	Condition	.03	.18	.17	.806
	Prejudice	.64	.09	7.30	<.001
	Prejudice × Condition	.67	.18	3.77	<.001

Note. Condition was coded 1 = Anti-Black, 0 = Anti-coworker. Overall models for expressive threat and freedom of speech were significant: $R^2 = .04$, $F(3, 344) = 5.37$, $p = .001$ and $R^2 = .16$, $F(3, 344 = 21.49)$, $p < .001$, respectively.
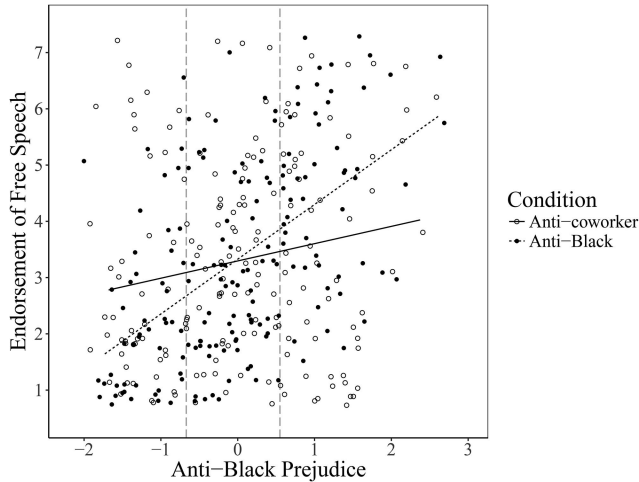
Figure 3. Anti-Black prejudice (standardized) more strongly predicted perceived violation of free speech in the anti-Black condition than in the anticoworker condition (Study 8). Participants that fall on the outside of the vertical lines differ by condition in their endorsement of freedom of speech.

anti-Black condition than the *anticoworker* condition, whereas participants scoring $Z = -0.67$ and lower in prejudice believed free speech was violated more in the *anticoworker* condition than in the *anti-Black* condition.

Expressive threat and perceived violation of free speech were correlated similarly in the *anti-Black*, $r = .32$, $p < .001$ and *anticoworker*, $r = .20$, $p = .007$ conditions, suggesting that a moderated mediation model might be possible.

Moderated mediation analysis. We examined if feeling greater threat to one's expression mediated the relationship between prejudice and perceived violation of free speech, but only in the *anti-Black* condition. We ran two separate linear regression models: first, we regressed expressive threat on prejudice (standardized), condition (mean-centered), and their interaction; second, we regressed free speech on expressive threat (mean-centered), prejudice (standardized), condition (mean-centered), and the prejudice by

condition interaction. Coefficients, indirect effects, and the index of moderated mediation (Hayes, 2015) were calculated using the PROCESS macro for SPSS (Model 8, Hayes, 2013) with 10,000 bias-corrected bootstrap resamples.

As reported in Table 1, the first regression equation yielded a significant prejudice by condition interaction on expressive threat, $p = .020$. The second regression equation yielded three significant predictors: expressive threat, $b = .28$, $SE = .07$, $t(343) = 3.78$, $p < .001$; prejudice, $b = .59$, $SE = .09$, $t(343) = 6.65$, $p < .001$; and the interaction, $b = .58$, $SE = .17$, $t(343) = 3.34$, $p = .001$.

The index of moderated mediation $= .08$ [.02, .21], indicating that the indirect effect in the *anti-Black* condition was significantly different from the indirect effect in the *anticoworker* condition (Hayes, 2015). We then examined the indirect effects by condition using the PROCESS macro (Model 4, Hayes, 2013), with 10,000 bias-corrected bootstrap resamples. As can be seen in Panels A and B of Figure 4, respectively, expressive threat mediated the relationship between prejudice and freedom of speech in the *anti-Black* condition, indirect effect $= .09$ [.02, .21]; however, the indirect effect was not significant in the *anticoworker* condition, indirect effect $= .02$ [$-.02$, .10].

Reading about an employee fired for racist speech made high prejudiced participants feel a threat to their autonomy; reading about an employee fired for insulting coworkers did not. The indirect effect of prejudice on freedom of speech through expressive threat was significant, but only when the circumstances suppressed anti-Black speech (i.e., the post that brought about the firing was anti-Black). Concerns about freedom, choice, and expressive threat are one of the reasons why prejudiced people justify another's prejudice. We have found correlationally *and* experimentally that expressive threat leads people to marshal free speech as a way to protect a sense of freedom to express their (socially unacceptable) views.

Threats to freedom elicit behaviors aiming to restore this loss of freedom (Brehm & Brehm, 1981; Wicklund, 1974). Antiprejudice social pressures that threaten freedom and evoke reactance are often pejoratively labeled as "political correctness" (e.g., Lalonde
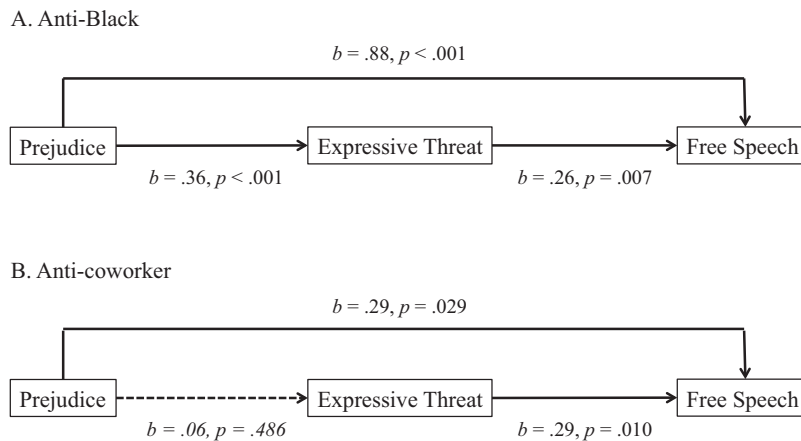
A. Anti-Black

B. Anti-coworker

Figure 4. Anti-Black prejudice (standardized) predicts greater expressive threat, which predicts greater perception that free speech has been violated. This indirect effect, however, is only present in the anti-Black condition (Study 8).

et al., 2000), and external pressures aimed at reducing prejudice may backfire (Brauer et al., 2012; Legault et al., 2011).

Small Scale Meta-Analysis

We calculated a meta-analytic effect size for the correlation between prejudice and freedom of speech in anti-Black conditions. Table 2 shows a meta-analytic r of .43 [.38, .48]; the meta-analytic r between prejudice and free speech in the three control conditions (in Studies 2, 3, and 8) was .09 [−.003, .19].

General Discussion

Prejudiced people justify another person's prejudiced speech. Across seven studies with 1,078 participants, anti-Black prejudice predicted how likely people were to claim that punishing someone for anti-Black prejudice violated their rights to freedom of speech. This relationship was reliably present in race-related contexts and reliably absent in race-neutral contexts, which suggests that the defense is not principled, but rather a justification and defense of the opportunity to express racist speech without substantial consequence.

People very *low* in prejudice demonstrated the opposite effect— they moved away from endorsing freedom of speech in racialized contexts. We did not predict this effect, but found it across three studies. It is unexpected (but we think persuasive) evidence that claiming free speech is a strategy with social implications that are understood—people low in overt prejudice avoid endorsing free speech values because they know it tends to justify and normalize racist speech.

Three studies provided no support for the *ego threat hypothesis,* that seeing another's expression of prejudice suppressed threatens a prejudiced person's self-integrity, how much they perceive themselves as a good and moral person. Three studies, using three proven self-relevant manipulations (self-affirmation, threat to unprejudiced self-image, similarity to protagonist), yielded no evidence to support the ego threat hypothesis. These studies had adequate power and nonsignificantly trended in the opposite direction of the hypothesis. Although these data do not eliminate the

Table 2
Meta-Analytic Relationship Between Prejudice and Freedom of Speech in Anti-Black Conditions

Sample	N	r
Study 1	176	.47 [.34, .59]
Study 2	126	.47 [.32, .60]
Study 3	125	.42 [.24, .57]
Study 4	206	.34 [.19, .47]
Study 5	134	.43 [.26, .59]
Study 6	135	.30 [.13, .45]
Study 8	176	.54 [.43, .64]
Meta-Analytic r and CI	1078	.43 [.38, .48]

Note. We calculated a fixed-effects effect size; all participants came from the same population and each of the constructs were measured similarly (Borenstein, Hedges, Higgins, & Rothstein, 2009). Cochran's Q was marginally significant, $Q(7) = 9.61$, $p = .062$, but calculating a random-effects meta-analytic effect size yielded the same effect size and an ever-so-slightly larger confidence interval, $r = .43$ [.37, 49].

hypothesis, they do not favor belief in it; support for the hypothesis is still notably absent.

By contrast, we found support for an *expressive threat hypothesis*— punishments for prejudice threaten the expressive freedom of prejudiced people, which leads them to use the value of free speech as a justification. The role of expressive threat appears only when participants were told about a man being fired for anti-Black prejudice (and not when he was fired for offensive speech about his coworkers)—the expressive threat concern was not principled, but rather protective of racial prejudice.

Justifying Another's Prejudice by Delegitimizing Suppressors

Although diminishing suppression has the same effect as enhancing justifications (more prejudice will be expressed), the underlying cause is the decrease in force used to inhibit prejudice expression and not an increase in the acceptability of prejudice (Crandall & Eshleman, 2003, p. 424).

Many prejudices are suppressed but prejudices (as affect) have motivational force toward expression. To reduce the conflict gracefully, people seek to justify their prejudice, to express them without negative sanction. The JSM is an individual model; it explains how people can to express their suppressed prejudices while avoiding sanctions. Our participants were justifying *another's* expressed prejudice, not their own, and the justification came *after* the expression, not before it.

Instead of offering a straightforward justification of racist speech, our participants protected their legitimacy to speak by delegitimizing the suppression (i.e., the firing), buffering the target from social disapproval and simultaneously protecting the self's right to expression. The claim of "free speech" can make disapproval of racist speech seem illegitimate, by characterizing it as a value violation (Fritsche, 2002; Weiner, Amirkhan, Folkes, & Verette, 1987). By inhibiting the suppressor, a prejudiced person can defend prejudiced speech which appears to be principled and content-neutral. Our data suggest this "pure" appeal to principle is in appearance only.

Why would third-parties frustrate the suppression of another's similarly prejudiced statements? Prejudiced people felt a greater threat to their expressive autonomy, which in turn motivated them to marshal freedom of speech as a justification for prejudice. In this way, the data are quite consistent with the general orientation of the JSM: People wish to express their attitudes, and will find ways to allow them to be expressed without guilt or public sanction. Justifications protect their freedom to communicate openly; prejudiced people defend not their self-integrity but rather their ability to let out their prejudices without sanction (see also Sittenthaler et al., 2016).

Values as Justifications

Values serve as effective justifications for political and intergroup attitudes because the values are "regarded as legitimate moral imperatives" (Kristiansen & Zanna, 1994, p. 50), and people can call upon them to defend a wide range of social attitudes (Kristiansen & Zanna, 1988, 1994). Americans see freedom of speech as an especially important value (Cohn, 2012; Wike & Simmons, 2015), which makes it an especially potent force to

undermine the legitimacy of prejudice suppression. Free speech values do not have a fundamental ideological connection to prejudice; in the absence of a specific racialized context, the relevance of free speech was unrelated to prejudiced attitudes.

Values reframe the situation (Lakoff, 2014); with the frame of "equality," firing someone for racist speech may be entirely acceptable, but with the frame of "free speech," the dismissal violates a cherished value. Functional theories of attitudes give values a value-expressive function (Ajzen, 2001; DeBono, 1987; Katz, 1960; Maio & Olson, 1994, 1995); expressing these attitudes are a concrete opportunity to articulate abstract beliefs. Inserting values as a frame serves an attitude-justifying function. Although values like freedom of speech might be a guiding principle to live by (Schwartz, 1992), appeals to speech values are not always principled.

Limits of Generality

We have shown, pretty reliably, that freedom of speech can be marshaled as a justification for another's punished prejudice. It seems that reactance motivates this effect, and we could not find evidence that threats to one's self motivates the effect. We take the opportunity to express what we think are the constraints on generality of our findings (Simons, Schoda, & Lindsay, 2017). The effect is likely to apply to expressing any prejudice that is counternormative—any "politically incorrect" prejudice that people might have should show this effect. Justifications are not interchangeable; they are tailored to the context, often with "narrow applicability" (Crandall & Eshleman, 2003). Because freedom of speech is a broadly endorsed value in the United States, it is likely to apply quite widely—any value that is both relevant and accepted by its audience should serve reasonably well. Our findings, however, are embedded in a culture with a deep appreciation of speech rights—other countries or cultures without this commitment will not show free speech effects. They may, however, find justification in other deeply held values when they are relevant to the situation.

Because the justification of another's contentious behavior can be costly, we suggest that these other-justifications will occur infrequently. However, they are likely to emerge when the situation prescribes them, when people are directly asked, when the observer has some personal responsibility for the expression or action, or when the observer has a very high level of the relevant prejudice. We also predict the converse of this; observers will distance themselves from—and avoid justifications of—the expressed prejudice to the extent they do not share the prejudice. This overt distancing should be emitted in the same instances that justifications will emerge for those who share the prejudice.

Avoiding Free Speech Endorsement

Our primary focus was on the justification of another's (suppressed) prejudice; our main hypotheses tested for equality of slopes—prejudice predicted more endorsement of freedom of speech in the racialized firing contexts than nonracialized ones. However, another way to probe these results is that participants scoring high in prejudice justified another's prejudice, whereas people scoring low in prejudice avoided defending another's right to freedom of speech when they were fired for racist speech.

We note, however, that the "side" of the prejudice distribution in which the effect of condition occurred is a function of the control condition, which had an arbitrary intercept. If we choose a control condition with highly normative speech, the intercept for "free speech endorsement" is very high, and most of the "significant" participants will be low in prejudice. However, if we choose a control statement that is much more unacceptable than the speech in the racialized condition, the control condition intercept will be low, and most of the "significant" participants will be those high in prejudice.

The original impetus for this research was the JSM, but this finding was not predicted because the JSM describes when people *have* prejudices, not when they lack them. Why do low prejudice participants disdain the justification? Presumably, they understand that endorsing free speech means justifying speech they disdain. These data suggest that highly prejudiced people do not have a monopoly on motivated reasoning.

In these "speech wars" debates (Millhiser, 2015), there is a tension between protecting people from prejudice and protecting freedom of speech. People high in prejudice protect free speech in a way that justifies prejudice, people low in prejudice avoid this justification in a way that fails to defend free speech rights. We focused on justifications for prejudice attributable to the social issues surrounding prejudice, and the intriguing notion of how people justify their abandonment of free speech rights goes unstudied in this paper.

Status of the Ego Threat Hypothesis

We tested the ego threat hypothesis using three different, previously proven methods. We find no support, but how confidently can we reject this hypothesis? The experiments have the advantage of being done in the context of several replications of the main hypothesis, where other manipulations were successful and reliable. We used three different operations to manipulate ego threat, offering the advantage of conceptual replication (Crandall & Sherman, 2016). Nevertheless, the experiments were conducted online; participants may be less threatened given the distance and impersonality of the situation. They may have more opportunity to resist the potency these kinds of manipulations. The online environment offers ecological validity because the phenomenon under study was mostly online in the first place, but in the online experimental situation, ego threats may seem less compelling.

Perhaps the ego threat is not to one's self-image, but rather to one's public image. Some theories of the self do not distinguish much between these two threats (e.g., Leary & Baumeister, 2000), but it is possible that there is some motivation to repair one's public image by defending another who's public image is threatened in the same way. We tested only the internal, self-image version of the hypothesis, and did not investigate self-presentation processes (Baumeister, 1982). There remains room to test the image-threat version of the hypothesis.

We have not completely ruled out the ego threat hypothesis. In Studies 4–6, the hypothesis fared poorly, despite using reasonable methods, adequate sample sizes, and in the context where other manipulations succeeded. We can say that, when describing this hypothesis to other psychologists, heads routinely nod vigorously when the ego threat hypothesis is stated; it is appealing. We suggest that the data described in Studies 4–6 should significantly reduce that appeal.

Conclusion

Dobby, a house-elf character from the *Harry Potter* book series, tells Harry Potter about the Room of Requirement: "It is a room that a person can only enter when they have real need of it. Sometimes it is there, and sometimes it is not, but when it appears, it is always equipped for the seeker's needs" (Rowling, 2003, pp. 386–387). Values, when used as justifications, operate in much the same way—they appear for the prejudiced person when it suits their needs but are absent when it does not. Freedom of speech is unprincipled, and therefore a justification because it only appears for prejudiced people when it is needed. Values may be used as guiding principles to live by, but they are also strategically deployed to justify prejudices.

References

Adams, G., Tormala, T. T., & O'Brien, L. T. (2006). The effect of self-affirmation on perception of racism. *Journal of Experimental Social Psychology, 42,* 616–626. http://dx.doi.org/10.1016/j.jesp.2005.11.001

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52,* 27–58. http://dx.doi.org/10.1146/annurev.psych.52.1.27

Allen, M. W., & Caillouet, R. H. (1994). Legitimation endeavors: Impression management strategies used by an organization in crisis. *Communication Monographs, 61,* 44–62. http://dx.doi.org/10.1080/03637759409376322

Andreoli, V. A., Worchel, S., & Folger, R. (1974). Implied threat to behavioral freedom. *Journal of Personality and Social Psychology, 30,* 765–771. http://dx.doi.org/10.1037/h0037529

Andreouli, E., Greenland, K., & Howarth, C. (2016). 'I don't think racism is that bad any more': Exploring the 'end of racism' discourse among students in English schools. *European Journal of Social Psychology, 46,* 171–184. http://dx.doi.org/10.1002/ejsp.2143

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology, 60,* 241–253. http://dx.doi.org/10.1037/0022-3514.60.2.241

Augoustinos, A., & Every, D. (2007). The language of "race" and prejudice: A discourse of denial, reason, and liberal-practical politics. *Journal of Language and Social Psychology, 26,* 123–141. http://dx.doi.org/10.1177/0261927X07300075

Baldwin, M., Biernat, M., & Landau, M. J. (2015). Remembering the real me: Nostalgia offers a window to the intrinsic self. *Journal of Personality and Social Psychology, 108,* 128–147. http://dx.doi.org/10.1037/a0038033

Bateman, D. (2015, September 9). *High school teacher fired after investigationinto 'racist' tweets*. Retrieved from http://www.thestar.com/yourtoronto/education/2015/09/09/high-school-teacher-fired-after-investigation-into-racist-tweets.html

Batson, C. D., Turk, C. L., Shaw, L. L., & Klein, T. R. (1995). Information function of empathic emotion: Learning that we value the other's welfare. *Journal of Personality and Social Psychology, 68,* 300–313. http://dx.doi.org/10.1037/0022-3514.68.2.300

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91,* 3–26. http://dx.doi.org/10.1037/0033-2909.91.1.3

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley. http://dx.doi.org/10.1002/9780470743386

Brauer, M., Er-rafiy, A., Kawakami, K., & Phills, C. E. (2012). Describing a group in positive terms reduces prejudices less effectively than describing it in positive and negative terms. *Journal of Experimental Social Psychology, 48,* 757–761. http://dx.doi.org/10.1016/j.jesp.2011.11.002

Brehm, J. W. (1966). *A theory of psychological reactance*. New York, NY: Academic Press.

Brehm, J. W. (1999). The intensity of emotion. *Personality and Social Psychology Review, 3,* 2–22. http://dx.doi.org/10.1207/s15327957pspr0301_1

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York, NY: Academic Press.

Case, B., & Weddington, T. (2015, November 12). *The "free speech" critiues of Missouri and Yale activists are hypocritical and racist*. Retrieved from http://www.wearepowershift.org/blogs/%E2%80%9Cfree-speech%E2%80%9D-critiques-missouri-and-yale-activists-are-hypocritical-and-racist

Chappell, B. (2015). *2 Oklahoma students seen in racist fraternity video apologize*. Retrieved from http://www.npr.org/sections/thetwo-way/2015/03/11/392279208/apologies-emerge-from-two-oklahoma-students-in-racist-sae-video

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology, 73,* 481–494. http://dx.doi.org/10.1037/0022-3514.73.3.481

Cobb, J. (2015, November 10). *Race and the free-speech diversion*. Retrieved from http://www.newyorker.com/news/news-desk/race-and-the-free-speech-diversion

Cohn, B. (2012, June 27). *21 charts that explain American values today*. Retrieved from http://www.theatlantic.com/national/archive/2012/06/21-charts-that-explain-american-values-today/258990/

Crandall, C. S. (1991). Multiple stigma and AIDS: Illness stigma and attitudes toward homosexuals and IV drug users in AIDS-related stigmatization. *Journal of Community & Applied Social Psychology, 1,* 165–172. http://dx.doi.org/10.1002/casp.2450010210

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin, 129,* 414–446. http://dx.doi.org/10.1037/0033-2909.129.3.414

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology, 82,* 359–378. http://dx.doi.org/10.1037/0022-3514.82.3.359

Crandall, C. S., Ferguson, M. A., & Bahns, A. J. (2013). When we see prejudice: The normative window and social change. In C. Stangor & C. Stangor (Eds.), *Stereotyping and prejudice* (pp. 53–69). New York, NY: Psychology Press.

Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology, 66,* 93–99. http://dx.doi.org/10.1016/j.jesp.2015.10.002

Crandall, C. S., Silvia, P. J., N'Gbala, A. N., Tsang, J. A., & Dawson, K. (2007). Balance theory, unit relations, and attribution: The underlying integrity of Heiderian theory. *Review of General Psychology, 11,* 12–30. http://dx.doi.org/10.1037/1089-2680.11.1.12

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108,* 593–623. http://dx.doi.org/10.1037/0033-295X.108.3.593

DeBono, K. G. (1987). Investigating the social-adjustive and value-expressive functions of attitudes: Implications for persuasion processes. *Journal of Personality and Social Psychology, 52,* 279–287. http://dx.doi.org/10.1037/0022-3514.52.2.279

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60,* 817–830. http://dx.doi.org/10.1037/0022-3514.60.6.817

Dijker, A. J. M. (1987). Emotional reactions to ethnic minorities. *European Journal of Social Psychology, 17,* 305–325. http://dx.doi.org/10.1002/ejsp.2420170306

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reaction. *Personality and Social Psy-*

chology Bulletin, 23, 316–326. http://dx.doi.org/10.1177/014616 7297233009

Dutton, D. G. (1971). Reactions of restaurateurs to Blacks and Whites violating restaurant dress requirements. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement, 3,* 298–302. http://dx.doi.org/10.1037/h0082272

Dutton, D. G. (1976). Tokenism, reverse discrimination, and egalitarianism in interracial behavior. *Journal of Social Issues, 32,* 93–107. http://dx .doi.org/10.1111/j.1540-4560.1976.tb02496.x

Farber, D. A. (1991). Free speech without romance: Public choice and the first amendment. *Harvard Law Review, 105,* 554–583. http://dx.doi.org/ 10.2307/1341698

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73,* 31–44. http://dx.doi.org/10.1037/0022-3514.73 .1.31

Fritsche, I. (2002). Account strategies for the violation of social norms: Integration and extension of sociological and social psychological typologies. *Journal for the Theory of Social Behaviour, 32,* 371–394. http://dx.doi.org/10.1111/1468-5914.00194

Frolik, C., & Gokavi, M. (2015, February 7). *2 Montgomery Co. deputies fired over text messages.* Retrieved from http://www.whio.com/news/ news/crime-law/results-of-deputies-racist-texts-investigation-to-/nj57h/

Gay, R. (2015, November 13). *The seduction of safety, on campus and beyond.* Retrieved from http://www.nytimes.com/2015/11/15/opinion/ sunday/the-seduction-of-safety-on-campus-and-beyond.html

Goldberg, M. (2015, November 17). *Is "free speech" becoming the new "all lives matter"?* Retrieved from http://www.slate.com/articles/double_x/ doublex/2015/11/free_speech_on_campus_will_it_become_the_new_all_ lives_matter.html

Haggerty, M. (2014, December 2). *"Racists Getting Fired" and the desire to do something about Ferguson.* Retrieved from http://www.onthemedia.org/ story/racists-getting-fired-and-desire-do-something-about-ferguson/

Hanson, H. (2015, May 14). *School principal fired after racist remark about 'Black people' at graduation.* Retrieved from http://www.huffingtonpost .com/2015/05/14/nancy-gordeuk-fired-racist-graduation_n_7283774.html

Harvey, R. D., & Oswald, D. L. (2000). Collective guilt and shame as m otivation for White support of Black programs. *Journal of Applied Social Psychology, 30,* 1790–1811. http://dx.doi.org/10.1111/j.1559-1816.2000.tb02468.x

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York, NY: Guilford Press.

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research, 50,* 1–22. http://dx.doi.org/10.1080/ 00273171.2014.962683

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106,* 766–794. http://dx.doi.org/10.1037/0033-295X.106.4.766

Henry, P. J., & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology, 23,* 253–283. http://dx.doi.org/10.1111/0162-895X.00281

Hensley, N. (2015, June 20). *Texas firefighter fired for siding with Charleston shooting suspect Dylann Roof in Facebook post.* Retrieved from http://www.nydailynews.com/news/national/texas-firefighter-fired-praising-dylan-roof-facebook-article-1.2265245

Heppner, W. L., Kernis, M. H., Nezlek, J. B., Foster, J., Lakey, C. E., & Goldman, B. M. (2008). Within-person relationships among daily self-esteem, need satisfaction, and authenticity. *Psychological Science, 19,* 1140–1145. http://dx.doi.org/10.1111/j.1467-9280.2008.02215.x

Hongo, H. (2015, June 30). *Citing First Amendment, Donald Trump sues Univision for $500 million.* Retrieved from http://gawker.com/citing-first-amendment-donald-trump-sues-univision-50-1715051971

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research and Politics, 2,* 1–12. http://dx.doi.org/10.1177/ 2053168015604648

Johnston, M. M., & Finney, S. J. (2010). Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the Basic Needs Satisfaction in General Scale. *Contemporary Educational Psychology, 35,* 280–296. http://dx.doi.org/10.1016/ j.cedpsych.2010.04.003

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly, 24,* 163–204. http://dx.doi.org/10.1086/266945

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. *Advances in Experimental Social Psychology, 38,* 283–357. http://dx.doi.org/10.1016/S0065-2601(06)38006-9

Kim, H. S., & Sherman, D. K. (2007). "Express yourself": Culture and the effect of self-expression on choice. *Journal of Personality and Social Psychology, 92,* 1–11. http://dx.doi.org/10.1037/0022-3514.92.1.1

Kimble, C. E., Kimble, E. A., & Croy, N. A. (1998). Development of self-handicapping tendencies. *The Journal of Social Psychology, 138,* 524–534. http://dx.doi.org/10.1080/00224549809600406

King, M. (2015). The 'knockout game': Moral panic and the politics of white victimhood. *Race & Class, 56,* 85–94. http://dx.doi.org/10.1177/ 0306396814567411

Klausner, A. (2015, November 20). *College student apologizes for criticizing Black lives matter online after her post got her fired from her job and has activists calling for her expulsion.* Retrieved from http://www .dailymail.co.uk/news/article-3327325/White-college-student-fired-restaurant-job-criticizing-Mizzou-Black-Lives-Matter-activists-online .html

Knowles, E. D., & Lowery, B. S. (2012). Meritocracy, self-concerns, and Whites' denial of racial inequity. *Self and Identity, 11,* 202–222. http:// dx.doi.org/10.1080/15298868.2010.542015

Knowles, E. D., Lowery, B. S., Chow, R. M., & Unzueta, M. M. (2014). Deny, distance, or dismantle? How White Americans manage a privileged identity. *Perspectives on Psychological Science, 9,* 594–609. http://dx.doi.org/10.1177/1745691614554658

Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology, 32,* 1134–1146. http://dx.doi.org/10.1037/0022-3514.32.6 .1134

Kristiansen, C. M., & Zanna, M. P. (1988). Justifying attitudes by appealing to values: A functional perspective. *British Journal of Social Psychology, 27,* 247–256. http://dx.doi.org/10.1111/j.2044-8309.1988 .tb00826.x

Kristiansen, C. M., & Zanna, M. P. (1994). The rhetorical use of values to justify social and intergroup values. *Journal of Social Issues, 50,* 47–65. http://dx.doi.org/10.1111/j.1540-4560.1994.tb01197.x

Lakoff, G. (2014). *The all new don't think of an elephant! Know your values and frame the debate.* White River Junction, VT: Chelsea Green Publishing.

Lalonde, R. N., Doan, L., & Patterson, L. A. (2000). Political correctness beliefs, threatened identifies, and social attitudes. *Group Processes & Intergroup Relations, 3,* 317–336. http://dx.doi.org/10.1177/1368 430200033006

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology, 32,* 1–62. http://dx.doi.org/10.1016/S0065-2601(00)80003-9

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of anti-prejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science, 22,* 1472–1477. http:// dx.doi.org/10.1177/0956797611427918

Litman, L., Robinson, J., & Atterbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral

sciences. *Behavior Research Methods*. Advance online publication. http://dx.doi.org/10.3758/s13428-016-0727-z

LoBianco, T., & Scott, E. (2015, November 12). *Trump blasts 'weak, ineffective' University of Missouri heads*. Retrieved from http://www.cnn.com/2015/11/12/politics/university-of-missouri-2016-candidates-react/

Lukianoff, G., & Haidt, J. (2015). *The coddling of the American mind*. Retrieved from http://www.theatlantic.com/magazine/archive/2015/09/the-coddling-of-the-american-mind/399356/

Maio, G. R., & Olson, J. M. (1994). Value-attitude-behavior relations: The moderating role of attitude functions. *British Journal of Social Psychology, 33,* 301–312. http://dx.doi.org/10.1111/j.2044-8309.1994.tb01027.x

Maio, G. R., & Olson, J. M. (1995). Relations betweenvalues, attitudes, and behavioral intentions: The moderating role of attitude function. *Journal of Experimental Social Psychology, 31,* 266–285. http://dx.doi.org/10.1006/jesp.1995.1013

Manne, K., & Stanley, J. (2015, November 13). *When free speech becomes a political weapon*. Retrieved from http://chronicle.com/article/When-Free-Speech-Becomes-a/234207/

McQueen, A., & Klein, W. P. (2006). Experimental manipulations of self-affirmation: A systematic review. *Self and Identity, 5,* 289–354. http://dx.doi.org/10.1080/15298860600805325

Millhiser, I. (2015, November 23). *How the campus speech wars could wind up in the Supreme Court*. Retrieved from http://thinkprogress.org/justice/2015/11/23/3725060/how-the-campus-speech-wars-could-wind-up-in-the-supreme-court/

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81,* 33–43. http://dx.doi.org/10.1037/0022-3514.81.1.33

Norman, B. (2015, July 27). *'It was meant to be funny,' says police officer fired for racist Facebook post*. Retrieved from http://www.local10.com/news/it-was-meant-to-be-funny-says-police-officer-fired-for-racist-facebook-post

O'Neill, B. (2014, November 22). *Free speech is so last century. Today's students want the 'right to be comfortable'*. Retrieved from http://new.spectator.co.uk/2014/11/free-speech-is-so-last-century-todays-students-want-the-right-to-be-comfortable/

Pearce, M. (2015, March 10). *Is University of Oklahoma frat's racist chant protected by 1st Amendment?* Retrieved from http://www.latimes.com/nation/la-na-oklahoma-fraternity-explainer-20150310-story.html

Phillips, L. T., & Lowery, B. S. (2015). The hard-knock life? Whites claim hardships in response to racial inequity. *Journal of Experimental Social Psychology, 61,* 12–18. http://dx.doi.org/10.1016/j.jesp.2015.06.008

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75,* 811–832. http://dx.doi.org/10.1037/0022-3514.75.3.811

Plant, E. A., & Devine, P. G. (2001). Responses to other-imposed pro-black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology, 37,* 486–501. http://dx.doi.org/10.1006/jesp.2001.1478

Plunkett, L. (2015, March 10). *Cop fired for making racist threats on Xbox Live*. Retrieved from http://kotaku.com/cop-fired-for-making-racist-threats-on-xbox-live-1690684050

Posner, R. A. (2002). Pragmatism versus purposivism in first amendment analysis. *Stanford Law Review, 54,* 737–752. http://dx.doi.org/10.2307/1229576

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185–227. http://dx.doi.org/10.1080/00273170701341316

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25,* 1–20.

Randazza, M. J. (2015, March 20). *What we risk when we ban racist speech*. Retrieved from http://www.cnn.com/2015/03/20/opinions/randazza-offensive-speech-first-amendment/

Reynolds, G. H. (2015, March 17). *OU broke the law to avoid bad press*. Retrieved from http://www.usatoday.com/story/opinion/2015/03/15/ou-oklahoma-frat-race-free-speech-column/24808377/

Rosenfeld, M. (2002). Hate speech in constitutional jurisprudence: A comparative analysis. *Cardozo Law Review, 24,* 1523–1556.

Rowling, J. K. (2003). *Harry Potter and the Order of the Phoenix*. New York, NY: Scholastic.

Rubin, M. (2015, November 11). *Demand students learn free speech*. Retrieved from https://www.commentarymagazine.com/american-society/college-students-free-speech/

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78. http://dx.doi.org/10.1037/110003-066X.55.1.68

Salancik, G. R., & Conway, M. (1975). Attitude inferences from salient and relevant cognitive content about behavior. *Journal of Personality and Social Psychology, 32,* 829–840. http://dx.doi.org/10.1037/0022-3514.32.5.829

Schlegel, R. J., Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology, 96,* 473–490. http://dx.doi.org/10.1037/a0014060

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25,* 1–65. http://dx.doi.org/10.1016/S0065-2601(08)60281-6

Shapiro, J. R., Mistler, S. A., & Neuberg, S. L. (2010). Threatened selves and differential prejudice expression by White and Black perceivers. *Journal of Experimental Social Psychology, 46,* 469–473. http://dx.doi.org/10.1016/j.jesp.2009.12.013

Sheldon, K. M., & Krieger, L. S. (2007). Understanding the negative effects of legal education on law students: A longitudinal test of self-determination theory. *Personality and Social Psychology Bulletin, 33,* 883–897. http://dx.doi.org/10.1177/0146167207301014

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology, 38,* 183–242. http://dx.doi.org/10.1016/S0065-2601(06)38004-5

Shipps, J. (2015, June 23). *Fired racist Alabama cop tries to cash in with support from Confederate group*. Retrieved from http://www.rawstory.com/2015/06/fired-racist-alabama-cop-tries-to-cash-in-with-support-from-confederate-group/

Sidanius, J., & Pratto, F. (1999). *Social dominance*. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139175043

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). *Constraints on generality (COG): A proposed addition to all empirical papers*. Unpublished manuscript. Retrieved from https://osf.io/phptw/

Sittenthaler, S., Jonas, E., & Traut-Mattausch, E. (2016). Explaining self and vicarious reactance: A process model approach. *Personality and Social Psychology Bulletin, 42,* 458–470. http://dx.doi.org/10.1177/0146167216634055

Sittenthaler, S., Traut-Mattausch, E., & Jonas, E. (2015). Observing the restriction of another person: Vicarious reactance and the role of self-construal and culture. *Frontiers in Psychology, 6,* 1052. http://dx.doi.org/10.3389/fpsyg.2015.01052

Skitka, L. J. (2009). Exploring the "lost and found" of justice theory and research. *Social Justice Research, 22,* 98–116. http://dx.doi.org/10.1007/s11211-009-0089-0

Smith, E. R., Coats, S., & Walling, D. (1999). Overlapping mental representations of self, in-group, and partner: Further response time evidence and a connectionist model. *Personality and Social Psychology Bulletin, 25,* 873–882. http://dx.doi.org/10.1177/0146167299025007009

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin, 104,* 23–35. http://dx.doi.org/10.1037/0033-2909.104.1.23

Snyder, C. R., Higgins, R. L., & Stucky, R. J. (1983). *Excuses: Masquerades in search of grace*. New York, NY: Wiley.

Soave, R. (2015, November 12). *Students: Fight racism, not free speech.* Retrieved from http://www.thedailybeast.com/articles/2015/11/12/students-fight-racism-not-free-speech.html

Spata, C. (2015, October 29). *Intern loses job after wildly racist tweet goes viral.* Retrieved from http://www.complex.com/pop-culture/2015/10/intern-loses-job-racist-tweet-cotton-n-word

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89,* 845–851. http://dx.doi.org/10.1037/0022-3514.89.6.845

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology, 21,* 261–302. http://dx.doi.org/10.1016/S0065-2601(08)60229-4

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology, 64,* 885–896. http://dx.doi.org/10.1037/0022-3514.64.6.885

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review, 4,* 290–299. http://dx.doi.org/10.1207/S15327957PSPR0404_1

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration, 23,* 263–280. http://dx.doi.org/10.1037/a0032359

Volokh, E. (2015, March 10). *No, it's not constitutional for the University of Oklahoma to expel students for racist speech.* Retrieved from https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/03/10/no-a-public-university-may-not-expel-students-for-racist-speech/

Weiner, B., Amirkhan, J., Folkes, V. S., & Verette, J. A. (1987). An attributional analysis of excuse giving: Studies of a naive theory of emotion. *Journal of Personality and Social Psychology, 52,* 316–324. http://dx.doi.org/10.1037/0022-3514.52.2.316

Wicklund, R. A. (1974). *Freedom and reactance*. Potomac, MD: Erlbaum.

Wike, R., & Simmons, K. (2015). *Global support for principle of free expression, but opposition to some forms of speech: Americans especially likely to embrace individual liberties*. Retrieved from Pew Research Center website: http://www.pewglobal.org/2015/11/18/global-support-for-principle-of-free-expression-but-opposition-to-some-forms-of-speech/

Worchel, S., & Arnold, S. E. (1973). The effects of censorship and attractiveness of the censor on attitude change. *Journal of Experimental Social Psychology, 9,* 365–377. http://dx.doi.org/10.1016/0022-1031(73)90072-3

Worchel, S., Arnold, S. E., & Baker, M. (1975). The effect of censorship on attitude change: The influence of censor and communication characteristics. *Journal of Applied Social Psychology, 5,* 227–239. http://dx.doi.org/10.1111/j.1559-1816.1975.tb00678.x

Workneh, L. (2015, May 27). *Ohio sheriff's deputy fired over racist tweets comparing Baltimore protesters to 'Planet of the Apes'.* Retrieved from http://www.huffingtonpost.com/2015/05/27/officer-racist-tweets-ape_n_7453458.html