Savage or satire: Individual differences in perceptions of disparaging and subversive racial humor

Stuart S. Miller*, Conor J. O'Dea, Tiffany J. Lawless, Donald A. Saucier

Kansas State University, United States of America

ARTICLE INFO

ABSTRACT

Subversive racial humor intends to confront and challenge racist beliefs to subvert traditional status hierarchies. However, because of its racial content, such humor can be misconstrued as disparaging. To understand why, we examined how individual differences relate to perceptions of disparaging and subversive racial humor. In three studies (Ns = 204, 134, and 200), college students and MTurk workers completed established measures of individual differences in tendencies to perceive racial prejudice (Studies 1–3), motivations to suppress racial prejudice (Study 3), and modern racism (Study 3), and responded to examples of disparaging or subversive racial humor. General tendencies to perceive racial prejudice were associated with more negative reactions to disparaging humor (e.g., perceiving it as less funny and more racist) and more positive reactions to subversive humor (e.g., perceiving it as more funny and more antiracist). Individual differences in motivations to suppress racial prejudice showed similar patterns, and modern racism showed opposite patterns. In general, our findings suggest that although subversive racial humor can be misconstrued, believing racial prejudice is pervasive and problematic is associated with a *greater* understanding and appreciation of subversive racial humor, an increasingly prevalent social phenomenon that requires further empirical attention.

1. Introduction

Racial humor is common in contemporary society, and has the ability to convey both antisocial and prosocial messages. Racial humor intended to reinforce status hierarchies (disparaging humor) increases individuals' endorsement of expressions of prejudice. Conversely, racial humor intended to subvert status hierarchies (subversive humor) has the potential to challenge existing status hierarchies by confronting people for their prejudice. However, individuals often interpret subversive racial humor as intending to disparage racial minorities, which may have unintended consequences.

Racial humor has the potential to reinforce as well as subvert status hierarchies between groups (Rappoport, 2005; Saucier, O'Dea, & Strain, 2016). For example, when individuals interpret a subversive racial joke as disparaging, they are more likely to endorse stereotypes about Black individuals (Saucier, Strain, Miller, O'Dea, & Till, 2018). Yet, there is no direct empirical evidence of the factors related to misattributing the intent of subversive racial humor. Our goal in the current studies it to examine the individual differences associated with interpretations of disparaging and subversive racial humor.

1.1. Disparaging racial humor

Rappoport (2005) describes disparaging humor as a sword meant to attack and marginalize lower-power social groups, reinforce existing status hierarchies, and normalize prejudice (e.g., Ferguson & Ford, 2008; Strain, Martens, & Saucier, 2016; Wicker, Barron III, & Willis, 1980). Disparaging humor targets individuals on the basis of uncontrollable (or difficult to control) qualities, such as race (e.g., Apte, 1987; Maio, Olson, & Bush, 1997; Weaver, 2010), sex (e.g., Ford, 2000; Ford, Boxer, Armstrong, & Edel, 2008; Ryan & Kanjorski, 1998; Thomas & Esses, 2004), and physical appearance (Baumeister & Carels, 2014). Although, blatant acts of prejudice are currently vilified, humor might provide a cover for expressing prejudice while avoiding negative social consequences (e.g., "it was just a joke"). Thus, the levity surrounding humor may increase the likelihood of a disparaging joke being told with impunity, and high-prejudiced individuals who recognize the potential to reinforce status hierarchies with little backlash may use humor as an outlet for expressing their socially unacceptable beliefs.

While exposure to disparaging humor generally loosens the norm condemning expressions of prejudice (Ford & Ferguson, 2004), confronting prejudice may strengthen this norm, reducing the likelihood

---

* Corresponding author at: Department of Psychological Sciences, Kansas State University, 492 Bluemont Hall, 1100 Mid-Campus Dr., Manhattan, KS 66506-5302, United States of America.
E-mail address: ssmiller@k-state.edu (S.S. Miller).

that individuals will express prejudice. However, this confrontation, while effective, comes at a cost. Specifically, the target of confrontation will likely resent the confronter (Czopp, Monteith, & Mark, 2006). Therefore, researchers have begun examining ways in which individuals may confront prejudice while avoiding this resentment. The perceived levity of humor may make it an effective method to confront prejudice while avoiding the social costs of confrontation (Saucier et al., 2016; Saucier et al., 2018).

1.2. Subversive racial humor

Rappoport (2005) describes racial humor used to confront prejudice as a shield to cope with adversity (Juni & Katz, 2001; Nezlek & Derks, 2001) and challenge racial prejudice (Boskin & Dorinson, 1985; Kramer, 2013). However, subversive humor can be problematic. For example, the majority of participants in a recent series of studies perceived subversive racial humor as being intended to disparage (Saucier et al., 2018). More concerningly, participants who misinterpreted the subversive humor were more likely to later endorse negative stereotypes about Black people—possibly because perceiving the humor as disparaging set the norm that disparaging beliefs are acceptable (Ford & Ferguson, 2004). However, when participants perceived the humor as intending to confront racial prejudices, they were more likely to see the humor as conveying the message that prejudice is not acceptable. Therefore, while confrontational forms of racial humor may function to subvert the social norms that reinforce racial animosity and enable expressions of prejudice, not everyone will get this message. And, although these findings suggest that subversive racial humor may be an effective way to confront prejudice, these results also suggest that the intent of subversive humor must be understood by the audience; otherwise efforts to use subversive humor as a shield could backfire.

1.3. Individual differences

The previous research suggests a potential problem with using subversive humor to confront prejudice, but it does not answer the question of what factors are related to how individuals perceive subversive racial humor differently. Theories of humor have identified many ways in which content may be perceived as funny. Jokes may be perceived as funny because they combine elements that are usually separate (incongruity theory; Deckers & Kizer, 1975), establish superiority over others (superiority theory; Zillmann & Cantor, 1976), and push the boundaries of appropriateness in surprising and tension-raising ways (relief theory; Shurcliff, 1968). It is important, though, for the experience of humor that the content not be perceived as excessively offensive (benign violation theory; McGraw & Warren, 2010), and that the content and intentions of the jokes are understood to be humorous (cognitive theory, Wyer & Collins, 1992). Given these conditions for the general experience of humor, there is good reason to assume that individual differences would be associated with perceptions of racial humor. Individuals' affective disposition toward, and identification with, the target of a joke (LaFave, 1972; Zillmann & Bryant, 1994), basic personality traits, such as neuroticism and extroversion (e.g., Galloway, 2010; Galloway & Chirico, 2008; Köhler & Ruch, 1996), as well as individual differences such as intelligence and strengths of the heart (Ruch, Heintz, Platt, Wagner, & Proyer, 2018), are associated with enjoyment of different types of humor. Individuals also differ in the extent to which they have cavalier attitudes about racist and sexist humor (Hodson, Rush, & MacInnis, 2010).

We extend the research on individual differences in perceptions of humor by examining how individuals' general tendencies to perceive prejudice are associated with the likelihood that individuals will interpret subversive and disparaging humor either as expressions of prejudice toward chronically disparaged groups (e.g., Black people), or as an effort to dismantle the thoughts and beliefs that support prejudice. In combination, individuals who more strongly believe racial prejudice

is pervasive, are more vigilant in looking for signs of prejudice, are more confident in being able to spot prejudice, and are less likely to trivialize the problems that racial prejudice causes, are more likely to perceive racism in a variety of contexts—from hypothetical scenarios, to real-world events, such as the shooting of Michael Brown in Ferguson, Missouri (Miller & Saucier, 2018). It is possible that tendencies to make attributions to prejudice may be related to perceiving subversive racial humor as disparaging. Because of their increased likelihood of making attributions to prejudice in ambiguous situations (Miller & Saucier, 2018), and because humor can be somewhat ambiguous as to its intended effects (e.g., is it just a joke meant to get a laugh, or is it something more sinister?), individuals with stronger tendencies to perceive prejudice may perceive racial humor as more prejudiced, regardless of the type of racial humor. Alternatively, such tendencies may be related to individuals' ability to tell the difference between disparaging and subversive racial humor. Research suggests higher tendencies to perceive prejudice are not associated with false alarms, such as perceiving expressions of prejudice in benign interracial interactions or in the absence of any evidence that prejudice is the cause of a behavior (Authors, unpublished data). We sought to test these competing hypotheses in the present studies.

We also sought to examine whether motivations to justify or suppress expressions of prejudice (Crandall & Eshleman, 2003) would be related to how individuals perceive disparaging and subversive racial humor. In general, we predicted greater motivations to justify prejudice would be associated with more favorable perceptions of disparaging racial humor, and that greater motivations to suppress prejudice would be associated with less favorable perceptions of disparaging racial humor. Individuals who are more intrinsically motivated to be free of racial prejudice, because it is an important part of their self-concept (Plant & Devine, 1998), are more likely to negatively evaluate disparaging racial humor because such motivations are negatively associated with lighthearted, nonchalant attitudes about humor that disparages low-status groups (Hodson et al., 2010). Perhaps, such appraisals would generalize to perceptions of subversive racial humor. Therefore, we examined whether greater internal motivations to suppress prejudice are associated with perceiving subversive racial humor as intending to disparage people of color or as intending to send a message that prejudice should not be tolerated.

On the other hand, individuals with greater levels of racial prejudice are more likely to construe disparaging racial humor as a non-serious attempt to get a laugh (Hodson et al., 2010), and thus be more likely to find such humor funny and downplay the racist message it conveys. If subversive racial humor is perceived as intending to express prejudice, then greater levels of prejudice should be related to seeing such humor as funny and nonracist. But, if subversive racial humor is perceived as intending to send an antiracist message, then greater levels of prejudice should be related to more negative evaluations of such humor.

1.4. Hypotheses

Our first hypothesis, the *generalization hypothesis*, predicts that higher tendencies to perceive prejudice, higher motivations to suppress prejudice, and lower levels of racially prejudiced attitudes would be associated with perceiving disparaging racial humor as more unfunny and racist, and that this association would generalize to perceiving subversive racial humor as similarly disparaging. Our second hypothesis is that individuals' expectations to see prejudice may be related to a deeper understanding of what is prejudiced and what is not. We labeled this hypothesis the *distinction hypothesis*. This hypothesis predicts the relationship between tendencies to perceive prejudice and perceptions of racial humor will be moderated by the type of racial humor being perceived. Specifically, higher tendencies to perceive prejudice, higher motivations to suppress prejudice, and lower levels of racially prejudiced attitudes will be associated with perceiving disparaging humor as prejudiced, but this relationship will be significantly weaker, or

potentially reversed in direction, when perceiving subversive humor.

In three studies, we measured participants' reactions to disparaging or subversive racial humor. We also measured tendencies to perceive prejudice (Studies 1–3), internal motivations to suppress prejudice (Study 3), and modern racism (Study 3) to test our hypotheses about how these individual differences would be associated with perceptions of both types of racial humor.

2. Study 1

In Study 1, we presented participants with either the subversive or disparaging joke used in Saucier et al. (2018) because the subversive joke used in their studies was often misinterpreted as disparaging. Additionally, we measured participants' levels of tendencies to perceive prejudice to see if this interacted with the type of racial humor in predicting participants' perceptions of the jokes. This allowed us to test our distinction hypothesis, which predicted an interaction, against the generalization hypothesis, which predicted no interaction.

2.1. Method

2.1.1. Participants

Participants (N = 204 American residents; 84% White; 55% male; ages 18 to 65, M = 34.93, SD = 11.11) were recruited using Amazon Mechanical Turk and were paid a small fee for participation. Our sample size was based on recent recommendations by Simmons, Nelson, and Simonsohn (2013) suggesting at least 50 participants per cell. We recruited more participants than that recommendation because we also tested the interaction between tendencies to perceive prejudice and joke condition. Demographic variables did not significantly differ by condition.

2.1.2. Tendencies to perceive prejudice

We measured participants' tendencies to perceive prejudice using the Propensity to Make Attributions to Prejudice scale (PMAPS). Previous research demonstrates the reliability and validity of the PMAPS (Miller & Saucier, 2018). The PMAPS contains 15 items measuring beliefs about the pervasiveness of racial prejudice (e.g., *You'll see lots of racism if you look for it*), trivialization of targets' concerns about racism (e.g., *Racial minorities are too worried about being discriminated against*), and vigilance (e.g., *I am on the lookout for instances of prejudice and discrimination*) and confidence in spotting racism (e.g., *I am quick to recognize prejudice*). Participants rated their agreement with the items on 1 (*Disagree Very Strongly*) to 9 (*Agree Very Strongly*) scales. We averaged all items (after reverse scoring the negatively worded items) to create composite scores where higher values represented greater tendencies to make attributions to prejudice (α = 0.90, M = 5.68, SD = 1.36; means and *alphas* were similar across the two conditions).

2.1.3. Procedure

Participants completed the PMAPS and then responded to one of two randomly assigned jokes in a between-participants design. In the subversive joke condition (n = 103), the joke was *What do you call a Black guy who flies a plane? A pilot you fucking racist!*; in the disparaging joke condition (n = 101) the joke was *Why do decent White people shop at Black people's yard sales? To get all their stuff back!* Participants rated how funny (*The joke I just read was funny*), offensive (*The joke I just read was offensive*), racist against Black people (*The joke I just read was racist against Black people*), racist against White people (*The joke I just read was racist against White people*), and anti-racist (*The joke I just read had an anti-racism message (i.e., racism is not okay)*) the joke was on 1 (*Disagree Very Strongly*) to 9 (*Agree Very Strongly*) scales. We treated these items as separate variables in our analyses.

Table 1
Perceptions of disparaging and subversive jokes (Study 1).

		b [95% CI]	PMAPS simple slopes	
			Disparaging [95% CI]	Subversive [95% CI]
Funny (R^2 = 0.12)	Disparaging – Subversive	−0.39 [−1.09, 0.30]		
	PMAPS		−0.92** [−1.28, −0.56]	0.17 [−0.20, 0.53]
	Joke × PMAPS ΔR^2 = 0.08	−1.09** [−1.61, −0.58]		
Offensive (R^2 = 0.30)	Disparaging – Subversive	2.22** [1.59, 2.85]		
	PMAPS		0.85** [0.52, 1.17]	0.55** [0.22, 0.88]
	Joke × PMAPS ΔR^2 = 0.01	0.30 [−0.17, 0.76]		
Racist against Blacks (R^2 = 0.48)	Disparaging – Subversive	4.07** [3.46, 4.69]		
	PMAPS		0.53** [0.21, 0.84]	0.22 [−0.11, 0.54]
	Joke × PMAPS ΔR^2 = 0.005	0.31 [−0.14, 0.76]		
Racist against Whites (R^2 = 0.10)	Disparaging – Subversive	−1.17** [−1.68, −0.65]		
	PMAPS		−0.20 [−0.47, 0.07]	−0.02 [−0.29, 0.25]
	Joke × PMAPS ΔR^2 = 0.004	−0.18 [−0.56, 0.20]		
Antiracist (R^2 = 0.50)	Disparaging – Subversive	−4.17** [−4.75, −3.59]		
	PMAPS		−0.25 [−0.53, 0.07]	0.20 [−0.10, 0.51]
	Joke × PMAPS ΔR^2 = 0.01	−0.43* [−0.86, −0.01]		

Note. PMAPS = propensity to make attributions to prejudice scale; regression coefficients are unstandardized.
* $p < .05$.
** $p < .01$.

2.2. Results and discussion

We entered PMAPS scores, joke condition, and their (mean centered) interaction as predictors in separate regression models for each of our measures of the perceptions of the jokes (see Table 1).[1] Consistent with our expectations, participants rated the disparaging joke as more offensive (Disparaging M = 6.49, SD = 2.34; Subversive M = 4.24, SD = 2.59; d = 0.91), more racist against Black people (Disparaging M = 7.84, SD = 1.86; Subversive M = 3.76, SD = 2.62; d = 1.80), less racist against White people (Disparaging M = 1.65, SD = 1.26; Subversive M = 2.83, SD = 2.30; d = 0.63), and less antiracist (Disparaging M = 1.71, SD = 1.46; Subversive M = 5.88, SD = 2.58; d = 1.99) than the subversive joke. There was no significant difference between the joke conditions on how funny participants perceived the joke to be (Disparaging M = 3.74, SD = 2.61; Subversive M = 4.15, SD = 2.74; d = 0.15).

Differences between the disparaging and subversive jokes in

---

[1] Zero-order correlations between all variables in Study 1 can be found in the online Supplementary materials.

perceptions of how funny the joke was, and the extent to which the joke contained an anti-racism message were moderated by PMAPS. Higher PMAPS scores were associated with perceiving the disparaging joke as less funny and more racist against Black people. PMAPS slopes for perceptions of the subversive joke as funny and racist supported the distinction hypothesis such that PMAPS scores were not significantly related to perceiving the subversive joke as funny or racist. Participants judged the subversive joke to be significantly more antiracist than the disparaging joke, but this difference became smaller at lower levels of PMAPS. Also, higher levels of PMAPS were associated with perceiving the subversive joke as funnier than the disparaging joke, but lower levels of PMAPS were associated with perceiving the disparaging joke as funnier than the subversive joke. These results suggest participants higher in PMAPS made a distinction between disparaging jokes and subversive jokes.

However, PMAPS slopes for perceptions of the subversive joke as offensive did support the generalization hypothesis: PMAPS scores were positively related to perceiving the subversive joke as offensive. These seemingly contradictory findings in both supporting and failing to support the generalization hypothesis could be due to how participants might have interpreted the item measuring perceived offensiveness because the subversive joke did contain language that could be construed as offensive (*you fucking racist!*). There are several reasons why someone might find offense in the subversive joke: (a) the joke refers to Black people, (b) the joke uses profanity, and (c) the joke makes fun of the listener. To rule out the first possibility, we conducted an exploratory analysis where we included perceptions of how racist against Black people the jokes were as a covariate in the regression model and observed a similar pattern of effects for the relationship between PMAPS scores and perceptions of offensiveness (Disparaging $b = 0.54$, $p < .001$; Subversive $b = 0.42$, $p = .002$). This shows that the relationship between tendencies to perceive prejudice and perceptions of offensiveness was independent of the extent to which participants perceived the jokes as racist against Black people.

In sum, our data suggest individuals who have greater tendencies to make attributions to prejudice are more likely to see disparagement humor as more racist, unfunny, and offensive. However, these tendencies to perceive prejudice do not appear to generalize to perceptions of subversive humor. We interpret these findings as providing stronger support for the distinction hypothesis than the generalization hypothesis.

3. Study 2

In Study 1, participants were only exposed to a single joke, limiting our ability to generalize our findings. Therefore, in Study 2, we exposed participants to a wider variety of disparaging and subversive racial humor using several Internet memes. Internet memes are common ways in which people experience racial humor (Authors, unpublished data) and often involve an image with superimposed text. We used a repeated measures design in Study 2 so participants could more directly compare the different forms of racial humor. We measured intended reactions to memes that are common online behaviors: share the meme by reposting it, "like" the meme, or confront the poster of the meme in the online comments. These behaviors are particularly interesting because they are typically non-anonymous and allow others to see a person's reaction to the meme.

3.1. Method

3.1.1. Participants

We recruited college students ($N = 134$; 80% White; 51% male; ages 18 to 33, $M = 19.52$, $SD = 2.42$) from introductory psychology classes. Sample size was based on recommendations of at least 50 people in a correlation analysis (see Wilson VanVoorhis & Morgan, 2007); we more than doubled that number because we planned to

compare the correlations between PMAPS and reactions to the memes across examples of disparaging and subversive types of humor.

3.1.2. Procedure

Participants first completed the PMAPS ($\alpha = 0.86$, $M = 5.61$, $SD = 1.06$). Next, using a within-participants design, participants responded (in a randomized order) to 24 disparaging memes about Black people, and 24 subversive memes (see Appendix A for examples). The research team, consisting of content-matter experts who have published on various forms of group humor, selected these memes from a pool gathered by their undergraduate research collaborators. The disparaging memes were selected because the researchers agreed the primary intention of the meme was to disparage Black people with negative stereotypes. The subversive memes were selected because the researchers agreed the primary intention of the meme was to subvert negative stereotypes about Black people. Memes in each set were selected to represent a broad range of well-known negative stereotypes about Black people and care was taken to make sure the memes were similar in text length and had the potential to be humorous, even if the content was offensive.[2]

Participants rated their agreement with nine statements measuring perceptions of how funny (*This meme is funny*), racist against Blacks (*This meme is racist toward Black people*), racist against Whites (*This meme is racist toward White people*), and anti-racist (*This meme has an anti-racism message (i.e., racism is not okay)*) each meme was. In addition, participants rated their agreement with items measuring their anticipated behavioral reaction to each meme: intentions to share the meme publicly (*I would share this meme publicly (e.g., Facebook, Twitter)*), intentions to share the meme privately (*I would share this meme privately (e.g., email a friend, show to someone I know, re-post it anonymously online)*), and intentions to confront someone online for posting the meme (*I would call someone out for sharing this online (e.g., in an online comment in response to its posting)*). We did not include the item to assess how offensive the meme was that we used in Study 1 because this item is vague, possibly assessing how much the meme is racist against Whites and/or Blacks, and/or how antiracist the meme is, and we preferred to assess those perceptions of the meme as offensive more specifically and directly with the respective items above. All items were measured on 1 (*Disagree Very Strongly*) to 9 (*Agree Very Strongly*) scales. We created composite variables by averaging each of the items across 24 disparaging memes and 24 subversive memes creating two composite variables for each of the 7 items above such that higher scores indicated greater agreement with each item across the different humor types (all αs > 0.87).

3.2. Results and discussion

Participants' perceptions of the memes significantly differed between the disparaging and subversive memes (see Table 2).[3] To test whether these differences in perceptions were moderated by PMAPS, we conducted correlations between PMAPS and reactions to the disparaging and subversive memes and tested the difference between these two sets of correlations (see Table 3). Higher levels of PMAPS were associated with rating the disparaging memes as less funny, more racist against Black people, and less antiracist. Higher levels of PMAPS were also associated with lesser intentions to share disparaging memes and greater intentions to confront someone who posted them. PMAPS was uncorrelated with perceptions of the disparaging joke as racist against Whites. These results indicate individuals higher in PMAPS show a greater understanding of the disparaging function of disparagement

---

[2] All materials for Study 2, can be found in the online Supplementary materials.

[3] Zero-order correlations between all variables in Study 2 can be found in the online Supplementary materials.

Table 2
Perceptions of disparaging and subversive memes (Study 2).

Measure	Disparaging memes M (SD)	Subversive memes M (SD)	M difference [95% CI]	η^2
Funny	2.96 (1.73)	3.34 (1.42)	−0.37** [−0.59, −0.16]	0.08
Racist against Blacks	6.87 (1.52)	3.99 (1.49)	2.87** [2.65, 3.10]	0.83
Racist against Whites	2.00 (1.00)	2.73 (1.18)	−0.73** [−0.85, −0.61]	0.52
Antiracist	1.74 (1.34)	3.92 (1.28)	−2.19** [−2.41, −1.96]	0.73
Share Publicly	1.59 (0.94)	1.98 (1.14)	−0.39** [−0.51, −0.27]	0.23
Share Privately	2.20 (1.54)	2.28 (1.32)	−0.09 [−0.25, 0.08]	0.01
Confront Online	2.95 (1.97)	2.09 (1.32)	0.86** [0.67, 1.05]	0.37

** $p < .01$.

Table 3
Correlations between PMAPS and perceptions of disparaging and subversive memes (Study 2).

Measure	Disparaging memes	Subversive memes	Difference [95% CI]
Funny	−0.27**	−0.01	−0.26 [−0.38, −0.13]
Racist against Blacks	0.51**	0.29**	0.23 [0.09, 0.37]
Racist against Whites	−0.07	−0.01	−0.06 [−0.16, 0.04]
Antiracist	−0.23**	−0.03	−0.20 [−0.37, −0.01]
Share Publicly	−0.19*	−0.04	−0.15 [−0.26, −0.03]
Share Privately	−0.25**	−0.12	−0.13 [−0.25, −0.02]
Confront Online	0.28**	0.05	0.21 [0.13, 0.32]

* $p < .05$.
** $p < .01$.

racial humor. Contrary to the generalization hypothesis, but supporting the distinction hypothesis, the pattern of results was not the same for subversive memes; PMAPS was uncorrelated with perceptions of the subversive memes as funny, racist against Whites, and antiracist. PMAPS was also uncorrelated with participants' intentions to share the meme publicly, privately, and to confront someone who shared the meme online. Higher levels of PMAPS were, however, positively related to perceptions of the subversive jokes as disparaging against Black individuals, but this correlation was not as strong as it was for disparaging memes.

Overall, PMAPS was more strongly correlated with reactions to disparaging memes than with reactions to subversive memes. These findings suggest levels of PMAPS moderate the differences in perceptions and reactions to subversive and disparaging humor. Thus, our results support the distinction hypothesis—which predicted this interaction—over the generalization hypothesis, which predicted PMAPS scores would be similarly related to reactions to disparaging and subversive humor.

4. Study 3

In Study 3, we included measures of internal motivations to suppress prejudice and modern racism to better understand the motivational components (justification and suppression factors) of reactions to racial humor. We also included a race-unrelated humor condition to test whether the effects were unique to racial humor or applied to humor more broadly.

4.1. Method

4.1.1. Participants

We recruited participants using Amazon Mechanical Turk and paid a small fee for participation. Because our justification and suppression

factor measures pertained to White people's beliefs, we removed from our analyses participants who were not White. The final sample ($N = 200$) consisted of White, American residents; 49% male; ages 18 to 98, $M = 35.36$, $SD = 12.11$. This sample size provided us more than the recommended 50 participants per cell (Simmons et al., 2013) in our design.

4.1.2. Justification and suppression factors

We used six items from the modern racism scale (MRS; McConahay, 1986) to measure participants' motivations to justify racial prejudice toward Black people (e.g., *Blacks are getting too demanding in their push for equal rights*). We used five items from the internal motivation scale (IMS; Plant & Devine, 1998) to measure participants' internal motivations to suppress prejudice (e.g., *Because of my personal values, I believe that using stereotypes about Black people is wrong*).[4] Participants responded to both measures on 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) scales. We averaged the items in each measure to create composite variables (MRS $\alpha = 0.90$, $M = 3.18$, $SD = 1.91$; IMS $\alpha = 0.86$, $M = 7.10$, $SD = 1.85$; means and *alphas* were similar across conditions) where higher scores represented higher levels of the construct.

4.1.3. Procedure

Participants completed the PMAPS ($\alpha = 0.91$, $M = 5.80$, $SD = 1.43$), the MRS, and the IMS in a randomized order. Next, participants were randomly assigned in a between-participants design to view either three disparaging, subversive, or race-unrelated neutral memes (see Appendix A). We selected three disparaging and three subversive memes from Study 2 that best portrayed the intent of that

---

[4] We also measured external motivations to suppress prejudice, but our primary predictions were about intrapsychic concerns and not about impression management concerns. We found that external motivations to suppress prejudice did not predict or moderate perceptions of the memes.

Table 4
Perceptions of disparaging, subversive, and neutral memes (Study 3).

	Neutral Memes M (SD)	Disparaging Memes M (SD)	Subversive Memes M (SD)	η²
Funny	6.47ᵃ (2.28)	3.69ᵇ (2.72)	4.33ᵇ (2.62)	0.19
Racist against Blacks	1.14ᶜ (0.60)	7.06ᵃ (2.50)	4.17ᵇ (2.83)	0.56
Racist against Whites	1.17ᵇ (0.70)	1.81ᵃ (1.58)	2.30ᵃ (1.78)	0.10
Antiracist	1.29ᶜ (0.94)	2.31ᵇ (2.16)	4.19ᵃ (2.81)	0.25
Share	5.01ᵃ (2.85)	2.01ᵇ (2.02)	2.67ᵇ (2.32)	0.23
Like	5.93ᵃ (2.75)	2.34ᵇ (2.53)	3.19ᵇ (2.55)	0.28
Confront	1.87ᵇ (1.58)	3.49ᵃ (2.27)	2.57ᵃᵇ (2.15)	0.09

Note. Results of one-way ANOVAs conducted on each item were all significant ($ps < 0.001$) Means in each row that do not share a superscript were significantly different at $p < .05$.

type of humor. The three neutral memes were sourced from the Internet and portrayed animals along with a humorous remark. After viewing the memes, participants rated their agreement, on 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) scales, with three separate items measuring different anticipated behavioral reactions to the memes (*I would share memes like these online* (e.g. *Twitter, Facebook*); *I would "like" a meme like these online* (e.g. *Twitter or Facebook "like" button*); *I would confront someone for sharing memes like these* (e.g. *in a comment in response to the post)*). Participants also rated, on 1 (*Not at All*) to 9 (*Very Much*) scales, how funny (*This type of humor is funny*), racist against Black people (*This type of humor is racist toward Black people*), racist against White people (*This type of humor is racist toward White people*), and antiracist (*This type of humor sends a message that racism is not okay*) the memes were. Each of these items were used as separate variables in our analyses.

4.2. Results and discussion

Table 4 displays the descriptive statistics and results of one-way ANOVAs on each of the items measuring reactions to the memes (all ANOVA $ps < 0.001$).[5] We used regression modeling techniques described in Study 1 to examine the moderators of perceptions of disparaging, subversive, and neutral memes. The meme condition was dummy coded and entered along with the continuous individual difference predictor and their (mean-centered) interactions as predictors of the different dependent measures of perceptions and anticipated behavioral reactions to the memes.

4.2.1. Tendencies to make attributions to prejudice

Table 5 shows the regression model results for PMAPS, the meme conditions, and their interactions. Except for responses about intentions to share the meme, PMAPS was unrelated to perceptions of the neutral memes, suggesting the associations between tendencies to make attributions to prejudice and perceptions of humor are specific to racial humor. Supporting the distinction hypothesis, but not the generalization hypothesis, how funny, racist against Blacks (but not Whites), and antiracist the different memes were perceived to be was moderated by PMAPS. Simple slopes showed tendencies to make attributions to prejudice were positively related to perceptions of how funny the subversive joke was, and negatively related to perceptions of how funny the disparaging joke was. Consistent with our findings in Study 1, perceptions of the disparaging memes as racist against Blacks were positively related to PMAPS, but PMAPS did not predict perceptions of the subversive joke as being racist against Black people. PMAPS did not predict perceptions of any of the memes as being racist against White

---
[5] Zero-order correlations between all variables in Study 3 can be found in the online Supplementary materials.

people. A unique finding of the current study was that PMAPS was positively related to perceiving the subversive memes as antiracist. PMAPS was not significantly related to perceiving the disparaging memes as antiracist. Higher PMAPS scores were associated with greater intentions to share and like the subversive memes, and lesser intentions to like the disparaging memes. Higher PMAPS scores were also related to greater intentions to confront a person who posts disparaging memes, but was unrelated to intentions to confront for subversive memes. These results support the distinction hypothesis and suggest individuals at higher levels of PMAPS perceive more of a difference between disparaging and subversive racial humor, and they are more able to identify the intended purpose of disparaging humor as an expression of racial prejudice toward minorities, and the intended purpose of subversive humor as an attempt to combat racial prejudice.

4.2.2. Modern racism

Table 6 shows the regression model results for the MRS, the meme conditions, and their interactions. Higher MRS scores were associated with perceiving the neutral memes as more funny, and with greater intentions to share and like the neutral memes, suggesting MRS might be tapping into something associated with perceptions of humor more generally. Consistent with the distinction hypothesis, MRS moderated perceptions of racism against Blacks, racism against Whites (marginally), and antiracism, as well as intentions to share and like the different memes. The positive relationship between MRS and perceptions of how funny the disparaging memes were supports our prediction that the justification of racial prejudice against Black people would be positively related to thinking disparaging humor is funny. MRS was negatively related to perceiving disparaging humor as racist against Blacks, but positively related to perceiving subversive humor as racist against Whites—suggesting individuals higher in MRS might understand the confrontational nature of subversive humor and see it as an affront to White people. Higher MRS scores were associated with higher levels of perceived antiracism for the disparaging memes and lower levels of perceived antiracism for the subversive memes—a pattern suggesting higher levels of MRS are associated with casting disparaging humor in a positive light while also disputing the prosocial message of subversive humor. As expected, greater intentions to share and like the disparaging memes were positively associated with MRS, but MRS did not predict intentions to confront. Together, these findings suggest individuals higher in modern racism have more positive perceptions of disparaging racial humor and more negative perceptions of subversive racial humor.

4.2.3. Internal motivations to suppress prejudice

Table 7 shows the regression model results for IMS, the meme conditions, and their interactions. IMS did not predict perceptions of neutral memes, suggesting the relationships between IMS and perceptions of humor are specific to racial humor. However, IMS was negatively related to intentions to share and like neutral memes, suggesting a relationship between IMS and behavioral reactions to humor in general. The distinction hypothesis was supported by the finding that IMS moderated the differences between memes in perceptions of racism against Black people, antiracism, and intentions to share and like the different memes. Consistent with the idea that IMS would be associated with negative perceptions of disparaging humor, higher levels of IMS were related to finding disparaging memes less funny, but IMS did not significantly predict perceptions of how funny the subversive memes were. There was a positive relationship between IMS and perceptions of subversive memes as racist against Black people, though this relationship was not as strong as for the disparaging memes, and the significant interaction supports the distinction hypothesis. However, we also interpret this finding as suggesting internal motivations to suppress prejudice might be related to "playing it safe" when it comes to judging the appropriateness racial humor, even if higher levels of IMS are associated with perceiving subversive humor as more antiracist. This conclusion that individuals at higher levels of IMS (compared to lower

Table 5

Perceptions of disparaging and subversive memes moderated by PMAPS (Study 3).

		b [95% CI]	Simple Slopes		
			Neutral [95% CI]	Disparaging [95% CI]	Subversive [95% CI]
Funny ($R^2 = 0.28$)	Neutral – Disparaging	2.89** [2.05, 3.67]			
	Subversive – Disparaging	0.78 [−0.05, 1.62]			
	Neutral – Subversive	2.08** [1.25, 2.91]			
	PMAPS		−0.26 [−0.63, 0.12]	−0.89** [−1.28, −0.56]	0.62** [0.20, 1.04]
Meme × PMAPS $\Delta R^2 = 0.09$, $F(2, 194) = 12.12$, $p < .001$					
Racist against Blacks ($R^2 = 0.60$)	Neutral –Disparaging	−5.94** [−6.65, −5.24]			
	Subversive – Disparaging	−2.94** [−3.67, −2.22]			
	Neutral – Subversive	−3.00** [−3.72, −2.28]			
	PMAPS		−0.002 [−0.33, 0.32]	0.87** [0.48, 1.25]	−0.18 [−0.54, 0.18]
Meme × PMAPS $\Delta R^2 = 0.04$, $F(2, 194) = 8.71$, $p < .001$					
Racist against Whites ($R^2 = 0.11$)	Neutral – Disparaging	−0.63* [−1.10, −0.15]			
	Subversive – Disparaging	0.46 [−0.03, 0.95]			
	Neutral – Subversive	−1.09** [−1.58, −0.60]			
	PMAPS		−0.02 [−0.24, 0.20]	−0.12 [−0.38, 0.14]	−0.20 [−0.45, 0.04]
Meme × PMAPS $\Delta R^2 = 0.01$, $F(2, 194) = 0.61$, $p = .546$					
Antiracist ($R^2 = 0.33$)	Neutral – Disparaging	−0.99** [−1.66, −0.32]			
	Subversive – Disparaging	2.02** [1.33, 2.71]			
	Neutral – Subversive	−3.01** [−3.70, −2.33]			
	PMAPS		−0.15 [−0.46, 0.15]	−0.36 [−0.73, 0.001]	0.77** [0.43, 1.12]
Meme × PMAPS $\Delta R^2 = 0.08$, $F(2, 194) = 11.80$, $p < .001$					
Share ($R^2 = 0.30$)	Neutral – Disparaging	3.09** [2.30, 3.88]			
	Subversive – Disparaging	0.79 [−0.02, 1.71]			
	Neutral – Subversive	2.30** [1.50, 3.10]			
	PMAPS		−0.41* [−0.77, −0.04]	−0.41 [−0.84, 0.02]	0.71** [0.30, 1.11]
Meme × PMAPS $\Delta R^2 = 0.07$, $F(2, 194) = 9.94$, $p < .001$					
Like ($R^2 = 0.34$)	Neutral – Disparaging	3.64** [2.81, 4.46]			
	Subversive – Disparaging	1.00* [0.16, 1.85]			
	Neutral – Subversive	2.64** [1.80, 3.47]			
	PMAPS		−0.20 [−0.58, 0.18]	−0.45* [−0.90, −0.002]	0.80** [0.38, 1.23]
Meme × PMAPS $\Delta R^2 = 0.06$, $F(2, 194) = 9.44$, $p < .001$					
Confront ($R^2 = 0.14$)	Neutral – Disparaging	−1.63** [−2.35, −0.90]			
	Subversive – Disparaging	−0.90* [−1.64, −0.15]			
	Neutral – Subversive	−0.73 [−1.47, 0.01]			
	PMAPS		−0.08 [−0.41, 0.26]	0.64** [0.25, 1.04]	0.25 [−0.12, 0.63]
Meme × PMAPS $\Delta R^2 = 0.03$, $F(2, 194) = 9.44$, $p = .026$					

Note. PMAPS = propensity to make attributions to prejudice scale; regression coefficients are unstandardized.

* $p < .05$.

** $p < .01$.

Table 6
Perceptions of disparaging and subversive memes moderated by MRS (Study 3).

		b [95% CI]	Simple Slopes		
			Neutral [95% CI]	Disparaging [95% CI]	Subversive [95% CI]
Funny ($R^2 = 0.26$)	Neutral – Disparaging	3.50** [2.91, 4.09]			
	Subversive – Disparaging	0.84 [−0.01, 1.69]			
	Neutral – Subversive	2.22** [1.38, 3.06]			
	MRS		0.32* [0.03, 0.62]	0.57** [0.28, 0.89]	0.07 [−0.30, 0.45]
Meme × MRS $\Delta R^2 = 0.02$, $F(2, 194) = 2.20$, $p = .114$					
Racist against Blacks ($R^2 = 0.60$)	Neutral –Disparaging	−6.08** [−6.80, −5.37]			
	Subversive – Disparaging	−3.06** [−3.78, −2.33]			
	Neutral – Subversive	−3.03** [−3.75, −2.31]			
	MRS		0.03 [−0.22, 0.28]	−0.54** [−0.79, −0.29]	0.09 [−0.23, 0.42]
Meme × MRS $\Delta R^2 = 0.03$, $F(2, 194) = 6.87$, $p = .001$					
Racist against Whites ($R^2 = 0.16$)	Neutral – Disparaging	−0.58* [−1.05, −0.11]			
	Subversive – Disparaging	0.56* [0.08, 1.04]			
	Neutral – Subversive	−1.14** [−1.61, −0.67]			
	MRS		0.04 [−0.12, 0.21]	0.34 [−0.02, 0.30]	0.35** [0.14, 0.56]
Meme × MRS $\Delta R^2 = 0.02$, $F(2, 194) = 2.52$, $p = .084$					
Antiracist ($R^2 = 0.32$)	Neutral – Disparaging	−0.87* [−1.56, −0.19]			
	Subversive – Disparaging	1.96** [1.26, 2.66]			
	Neutral – Subversive	−2.84** [−3.53, −2.15]			
	MRS		0.15 [−0.09, 039]	0.35** [0.11, 0.59]	−0.49** [−0.80, −0.18]
Meme × MRS $\Delta R^2 = 0.07$, $F(2, 194) = 9.27$, $p < .001$					
Share ($R^2 = 0.28$)	Neutral – Disparaging	3.22** [2.41, 4.02]			
	Subversive – Disparaging	0.76 [−0.06, 1.58]			
	Neutral – Subversive	2.46** [1.64, 3.27]			
	MRS		0.38* [0.10, 0.66]	0.36* [0.08, 0.64]	−0.19 [−0.55, 0.18]
Meme × MRS $\Delta R^2 = 0.03$, $F(2, 194) = 3.52$, $p = .032$					
Like ($R^2 = 0.33$)	Neutral – Disparaging	3.81** [2.97, 4.64]			
	Subversive – Disparaging	0.96* [0.11, 1.82]			
	Neutral – Subversive	2.85** [2.00, 3.69]			
	MRS		0.35* [0.06, 0.65]	0.40** [0.11, 0.69]	−0.27 [−0.64, 0.11]
Meme × MRS $\Delta R^2 = 0.03$, $F(2, 194) = 4.35$, $p = .014$					
Confront ($R^2 = 0.10$)	Neutral – Disparaging	−1.67** [−2.41, −0.92]			
	Subversive – Disparaging	−0.98* [−1.75, −0.22]			
	Neutral – Subversive	−0.68 [−1.44, 0.07]			
	MRS		0.01 [−0.25, 0.28]	−0.15 [−0.41, 0.11]	−0.21 [−0.55, 0.13]
Meme × MRS $\Delta R^2 = 0.01$, $F(2, 194) = 0.62$, $p = .538$					

Note. MRS = modern racism scale; regression coefficients are unstandardized.
* $p < .05$.
** $p < .01$.

Table 7
Perceptions of disparaging and subversive memes moderated by IMS (Study 3).

		b [95% CI]	Simple Slopes		
			Neutral [95% CI]	Disparaging [95% CI]	Subversive [95% CI]
Funny ($R^2 = 0.25$)	Neutral – Disparaging	2.96** [2.13, 3.79]			
	Subversive – Disparaging	0.78 [−0.07, 1.63]			
	Neutral – Subversive	2.18** [1.34, 3.02]			
	IMS		0.20 [−0.53, 0.12]	−0.59** [−0.90, −0.27]	−0.26 [−0.59, 0.06]
Meme × IMS $\Delta R^2 = 0.01$, $F(2, 194) = 2.20$, $p = .192$					
Racist against Blacks ($R^2 = 0.64$)	Neutral –Disparaging	−6.08** [−6.76, −5.41]			
	Subversive – Disparaging	−3.06** [−3.75, −2.37]			
	Neutral – Subversive	−3.02** [−3.70, −2.34]			
	IMS		−0.06 [−0.32, 0.21]	0.78** [0.53, 1.04]	0.31* [0.04, 0.57]
Meme × IMS $\Delta R^2 = 0.04$, $F(2, 194) = 10.34$, $p < .001$					
Racist against Whites ($R^2 = 0.10$)	Neutral – Disparaging	0.61* [−1.09, −0.13]			
	Subversive – Disparaging	0.51* [0.01, 1.00]			
	Neutral – Subversive	−1.11** [−1.60, −0.63]			
	IMS		−0.07 [−0.26, 0.12]	−0.05 [−0.23, 0.14]	−0.01 [−0.20, 0.18]
Meme × IMS $\Delta R^2 = 0.001$, $F(2, 194) = 0.12$, $p = .886$					
Antiracist ($R^2 = 0.34$)	Neutral – Disparaging	−0.91** [−1.59, −0.24]			
	Subversive – Disparaging	1.94** [1.25, 2.63]			
	Neutral – Subversive	−2.86** [−3.54, −2.18]			
	IMS		−0.23 [−0.49, 0.03]	−0.28* [−0.54, −0.03]	0.56** [0.29, 0.82]
Meme × IMS $\Delta R^2 = 0.09$, $F(2, 194) = 12.54$, $p < .001$					
Share ($R^2 = 0.30$)	Neutral – Disparaging	3.21** [2.42, 4.00]			
	Subversive – Disparaging	0.73 [−0.08, 1.54]			
	Neutral – Subversive	2.47** [1.67, 3.28]			
	IMS		−0.59** [−0.90, −0.28]	−0.36* [−0.66, −0.06]	0.11 [−0.20, 0.42]
Meme × IMS $\Delta R^2 = 0.03$, $F(2, 194) = 3.52$, $p = .032$					
Like ($R^2 = 0.32$)	Neutral – Disparaging	3.75** [2.91, 4.59]			
	Subversive – Disparaging	0.94* [0.08, 1.80]			
	Neutral – Subversive	2.81** [1.96, 3.66]			
	IMS		−0.34* [−0.67, −0.01]	−0.40* [−0.72, −0.08]	0.21 [−0.12, 0.54]
Meme × IMS $\Delta R^2 = 0.03$, $F(2, 194) = 4.13$, $p = .017$					
Confront ($R^2 = 0.12$)	Neutral – Disparaging	−1.67** [−2.41, −0.93]			
	Subversive – Disparaging	−0.99* [−1.75, −0.24]			
	Neutral – Subversive	−0.68 [−1.43, 0.07]			
	IMS		−0.09 [−0.38, 0.20]	0.32* [0.04, 0.60]	0.16 [−0.13, 0.45]
Meme × IMS $\Delta R^2 = 0.02$, $F(2, 194) = 2.04$, $p = .133$					

Note. IMS = internal motivations scale; regression coefficients are unstandardized.
* $p < .05$.
** $p < .01$.

levels of IMS) might be more hesitant to say subversive humor is not racist, while at the same time, showing a greater understanding of the intent of subversive humor, was also supported by the finding that higher IMS scores were associated with more strongly agreeing that subversive memes were antiracist. IMS was also negatively related to perceiving the disparaging memes as antiracist. Also supporting the distinction hypothesis, higher IMS scores were associated with lower intentions to share and like the disparaging memes, and greater intentions to confront someone who posted these memes, but IMS was not related to behavioral intentions regarding the subversive memes. These findings support the distinction hypothesis that greater internal motivations to suppress prejudice would be associated with perceiving more of a difference between disparaging versus subversive forms of humor. More specifically, these findings support the idea that IMS is associated with understanding the intent of disparaging humor to reinforce status hierarchies, and the intent of subversive humor to challenge status hierarchies.

4.2.4. Conclusions

The majority of the results in Study 3 support the distinction hypothesis; little support for the generalization hypothesis was found. These finding suggest greater tendencies to perceive prejudice, greater motivations to suppress prejudice, and lesser endorsement of justifications for racial prejudice are associated with differential reactions to disparaging and subversive forms of humor. These results also suggest there may be a motivational component to how racial humor is perceived. Individuals higher in modern racism—who have a greater need to justify their prejudicial beliefs—may be more motivated to deny disparaging humor is racist against Black people, deny subversive humor contains an antiracist message, and claim subversive humor is racist against White people. The role of motivations to suppress expressions of prejudice appears to be more complicated, however. While higher levels of IMS were associated with perceiving disparaging humor to be more racist against Black people, it was also (although more weakly) associated with perceiving subversive humor as more racist against Black people. Yet, higher levels of IMS were associated with being more likely to see the antiracist message of subversive memes. The combination of these last two findings suggests that while individuals with greater motivations to suppress prejudice might be more likely to perceive the antiracist intent of subversive humor, they may also be somewhat reluctant to say that subversive humor is not racist against Black people. In sum, Study 3 illustrates how individual differences impact perceptions and reported understanding of disparaging and subversive humor.

5. General discussion

Racial humor has the potential to be used as a sword that reinforces prejudice or a shield that challenges racist beliefs (Rappoport, 2005; Saucier et al., 2016). However, humor can send ambiguous messages, and thus there is the potential that subversive humor could be interpreted as disparaging. The current research elucidates how individual differences relate to perceptions of subversive and disparaging racial humor. Although many people interpret subversive racial humor as disparaging (Saucier et al., 2018), the current studies do not suggest this misunderstanding is due to greater tendencies to make attributions to prejudice.

In all three studies, higher levels of PMAPS were associated with perceiving disparaging racial humor as more racist against Black people than subversive racial humor. Additionally, higher levels of PMAPS were associated with perceiving subversive racial humor as funnier and more antiracist than disparaging racial humor. Higher levels of PMAPS were also associated with being more willing to share and "like" subversive racial humor online. These findings support the distinction hypothesis that tendencies to make attributions to prejudice moderate different perceptions of the two forms of racial humor. While greater

tendencies to make attributions to prejudice are associated with perceiving disparaging racial humor as racist, we found little evidence that this generalizes to perceiving subversive racial humor as prejudiced. Rather, our results suggest greater tendencies to make attributions to prejudice are related to perceiving subversive racial humor as intended to challenge racial prejudice.

In Study 3, we also examined how motivations to justify or suppress racial prejudice were associated with perceptions of racial humor. Consistent with the hypothesis that greater levels of modern racism are associated with greater tendencies to justify or downplay expressions of racial prejudice (Crandall & Eshleman, 2003), modern racism predicted more favorable reactions to disparaging humor (e.g., it's not racist, it's funny). Additionally, greater levels of modern racism were associated with more unfavorable reactions to subversive humor (e.g., it's racist against Whites), consistent with findings that modern racism is associated with negative attitudes about programs that are aimed at increasing racial equality (e.g., Awad, Cokley, & Ravitch, 2005). Such evidence suggests modern racism is associated with motivations to undermine attempts to challenge racial prejudice. Interestingly, modern racism was associated with perceiving subversive memes as more racist against Whites, but not more racist against Blacks. Thus, at higher levels of modern racism, individuals are not interpreting subversive humor as making fun of Black people—suggesting they are not misunderstanding the intent of subversive racial humor.

Like tendencies to perceive prejudice, greater motivations to suppress prejudice were associated with perceiving subversive racial humor as funnier and more antiracist than disparaging racial humor (Study 3). But, unlike tendencies to perceive prejudice, greater levels of motivations to suppress prejudice were also associated with perceiving subversive humor as more racist against Blacks. That PMAPS was unrelated to perceiving subversive racial humor as racist against Blacks, but IMS was, might imply IMS is tapping into a motivational component associated with perceiving prejudice not captured by PMAPS. Individuals who are higher in IMS might be concerned with making a mistake, and would rather err on the side of caution (see Barrett & Swim, 1998) when making judgements about racial humor, even if they see a potentially antiracist message. Saying something was not racist when it was might be a more serious error than saying something was racist when it was not.

Alternatively, some individuals who are more internally motivated to suppress prejudice might be interpreting subversive humor as conveying a very different message—one that indirectly reinforces the existence of stereotypes and prejudice. The racial memes used in Study 3 had a setup that primed people to think stereotypically (e.g., "What's black and never works?") and then confronts the perceiver for potentially thinking stereotypically ("Decaffeinated coffee you racist bastard!"). One could argue that making fun of how common it is for people to think stereotypically perhaps normalizes the existence of prejudice and stereotyping. If this were the case, then subversive racial humor would have similar effects as disparaging humor in loosening the norms that work to suppress expressions of prejudice. Clearly, this is a direction for future research.

Future research on racial humor should also consider the potential interactions between preferences for different humor styles and personality traits (e.g., Ruch et al., 2018), and identification with the targets (e.g., LaFave, 1972) of racial humor, in predicting reactions to racial humor. Similarly, research could examine how these individual differences interact with the context in which subversive racial humor is presented to affect how such humor is perceived.

5.1. Limitations

One caveat that limits our ability to generalize our findings is that we used a convenience sample of college students in Studies 1 and 2, and even though we used an MTurk sample in Study 3, we did not have true community samples. Our samples also included few people of color

which prevented us from examining (with any confidence) whether perceptions of different forms of racial humor are moderated by racial identity. Differences in age (e.g., Ruch, McGhee, & Hehl, 1990) and racial identity are likely to affect how people perceive racial humor.

Another limitation of the current studies is that we could not fully represent the range of subversive and disparaging racial humor, and our conclusions must be constrained to the examples we chose as stimuli in our studies. Our results showed subversive memes were perceived to be more funny and less racist than disparaging memes on average, but there was variability within our examples of each type (especially in Study 2 in which participants viewed 24 examples of each type). Our studies were not designed to examine how the content of different types of racial humor (e.g., jokes making fun of stereotypes versus racist behavior) are perceived. Both the content and structure of a joke affect how it is perceived (Carretero-Dios, Pérez, & Buela-Casal, 2010; Ruch, 1992). We did not account for the possible variations in content or structure that might have contributed to the differences we found in perceptions of disparaging and subversive racial humor. This is a clear direction for future research. For example, even disparaging jokes could be perceived as being cleverly constructed and finding such jokes funny could affect how racist they are perceived to be. If an initial reaction is to laugh, only moments later to realize the disparaging message of the joke, people may be motivated to downplay how racist the joke was to protect their self-image as being non-racist. However, the initial light-hearted reaction could set the norm that racially prejudiced attitudes are acceptable (Ford & Ferguson, 2004). Similarly, the structure of

some forms of subversive humor might interfere with individuals' ability to get the intended antiracist message—causing the joke to be misconstrued and possibly backfire by weakening, rather than strengthening, the norms condemning expressions of prejudice.

5.2. Implications and conclusion

Humor can be powerful. It can take people down or build them up. But unfortunately, humor can be misinterpreted. Thus, using humor to send an antiracist message is a risky strategy. We found that individuals' understanding of forms of racial humor is related to individual differences related to beliefs and motivations about expressions of prejudice. Individuals higher in modern racism may be motivated to downplay the severity of disparaging racial humor and negatively evaluate subversive racial humor. Conversely, individuals higher in motivations to suppress prejudice may be more likely to recognize and appreciate the antiracist sentiment of subversive racial humor, but also be reluctant to say such humor as not racist. And, for those who are more likely to see racism, it appears they are more likely to distinguish when humor is being used to disparage racial outgroups versus to subvert racial hierarchies. In short, racial humor, despite its intention, is open to various interpretations. Even racial humor intended to bring attention to and combat continued expressions of racism in contemporary society may be perceived as sword or shield, as reinforcing or subverting racial hierarchies, as perpetuating or confronting expressions of prejudice, as savage or satire.

Appendix A. Study 3 memes

Subversive memes

Disparaging memes

Neutral memes

Appendix B. Supplementary data and materials

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2019.01.029.

References

Apte, M. L. (1987). Ethnic humor versus "sense of humor". *The American Behavioral Scientist, 30*, 27–41. https://doi.org/10.1177/000276487030003004.

Awad, G. H., Cokley, K., & Ravitch, J. (2005). Attitudes toward affirmative action: A comparison of color-blind versus modern racist attitudes. *Journal of Applied Social Psychology, 35*, 1384–1399. https://doi.org/10.1111/j.1559-1816.2005.tb02175.x.

Barrett, L., & Swim, J. K. (1998). Appraisals of prejudice and discrimination. In J. K. Swim, & C. Stangor (Eds.). *Prejudice: The target's perspective* (pp. 11–36). San Diego, CA: Academic.

Baumeister, J. M., & Carels, R. A. (2014). Weight-related humor in the media: Appreciation, distaste and anti-fat attitudes. *Psychology of Popular Media Culture, 3*, 223–238. https://doi.org/10.1037/2376-6972.1.S.92.

Boskin, J., & Dorinson, J. (1985). Ethnic humor: Subversion and survival. *American Quarterly, 37*, 81–97. https://doi.org/10.2307/2712764.

Carretero-Dios, H., Pérez, C., & Buela-Casal, G. (2010). Assessing the appreciation of the content and structure of humor: Construction of a new scale. *Humor, 23*, 307–325. https://doi.org/10.1515/humr.2010.014.

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

Czopp, A. M., Monteith, M. J., & Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of Personality and Social Psychology, 90*, 784–803. https://doi.org/10.1037/0022-3514.90.5.784.

Deckers, L., & Kizer, P. (1975). Humor and the incongruity hypothesis. *The Journal of Psychology, 90*, 215–218. https://doi.org/10.1080/00223980.1975.9915778.

Ferguson, M. A., & Ford, T. E. (2008). Disparagement humor: A theoretical and empirical review of psychoanalytic superiority, and social identity theories. *Humor, 21*, 283–312. https://doi.org/10.1515/HUMOR.2008.014.

Ford, T. E. (2000). Effects of sexist humor on tolerance of sexist events. *Personality and Social Psychology Bulletin, 26*, 1094–1107. https://doi.org/10.1177/0146167200261106.

Ford, T. E., Boxer, C. F., Armstrong, J., & Edel, J. R. (2008). More than "just a joke": The prejudice-releasing function of sexist humor. *Personality and Social Psychology Bulletin, 34*, 159–170. https://doi.org/10.1177/0146167207310022.

Ford, T. E., & Ferguson, M. A. (2004). Social consequences of disparagement humor: A prejudiced norm theory. *Personality and Social Psychology Review, 8*, 79–94. https://doi.org/10.1207/s15327957pspr0801_4.

Galloway, G. (2010). Individual differences in personal humor styles: Identification of prominent patterns and their associates. *Personality and Individual Differences, 48*, 563–567. https://doi.org/10.1016/j.paid.2009.12.007.

Galloway, G., & Chirico, D. (2008). Personality and humor appreciation: Evidence of an association between trait neuroticism and preferences for structural features of humor. *Humor, 21*, 129–142. https://doi.org/10.1515/HUMOR.2008.006.

Hodson, G., Rush, J., & MacInnis, C. C. (2010). A joke is just a joke (except when it isn't): Cavalier humor beliefs facilitate the expression of group dominance motives. *Journal of Personality and Social Psychology, 99*, 660–682. https://doi.org/10.1037/a0022671.

Juni, S., & Katz, B. (2001). Self-effacing wit as a response to oppression: Dynamics in ethnic humor. *The Journal of General Psychology, 128*, 119–142. https://doi.org/10.1080/00221300109598903.

Köhler, G., & Ruch, W. (1996). Sources of variance in current sense of humor inventories: How much substance, how much method variance? *Humor, 9*, 363–397. https://doi.org/10.1515/humr.1996.9.3-4.363.

Kramer, C. A. (2013). An existentialist account of the role of humor against oppression. *Humor, 26*, 629–651. https://doi.org/10.1515/humor-2013-0045.

LaFave, L. (1972). Humor judgments as a function of reference groups and identification classes. In J. H. Goldstein, & P. E. McGhee (Eds.). *The psychology of humor: Theoretical perspectives and empirical issues* (pp. 195–210). New York, NY: Academic Press.

Maio, G. R., Olson, J. M., & Bush, J. E. (1997). Telling jokes that disparage social groups: Effects on the joke teller's stereotypes. *Journal of Applied Social Psychology, 27*, 1986–2000. https://doi.org/10.1111/j.1559-1816.1997.tb01636.x.

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio, & S. L. Gaertner (Eds.). *Prejudice, discrimination, and racism* (pp. 91–125). New York, NY: Academic Press.

McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological Science, 21*, 1141–1149. https://doi.org/10.1177/0956797610376073.

Miller, S. M., & Saucier, D. A. (2018). Individual differences in the propensity to make attributions to prejudice. *Group Processes & Intergroup Relations, 21*, 280–301. https://doi.org/10.1177/1368430216674342.

Nezlek, J. B., & Derks, P. (2001). Use of humor as a coping mechanism, psychological adjustment, and social interaction. *Humor, 14*, 395–413. https://doi.org/10.1515/humr.2001.011.

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

Rappoport, L. (2005). *Punchlines: The case for racial, ethnic, and gender humor.* Connecticut: Praeger Publishers.

Ruch, W. (1992). Assessment of appreciation of humor: Studies with the 3 WD humor test. In C. D. Speilberger, & J. N. Butcher (Eds.). *Advances in personality assessment* (pp. 27–75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ruch, W., Heintz, S., Platt, T., Wagner, L., & Proyer, R. T. (2018). Broadening humor: Comic styles differentially tap into temperament, character, and ability. *Frontiers in Psychology, 9*, 6. https://doi.org/10.3389/fpsyg.2018.00006.

Ruch, W., McGhee, P. E., & Hehl, F. J. (1990). Age differences in the enjoyment of incongruity-resolution and nonsense humor during adulthood. *Psychology and Aging, 5*, 348–355. https://doi.org/10.1037/0882-7974.5.3.348.

Ryan, K. M., & Kanjorski, J. (1998). The enjoyment of sexist humor, rape attitudes, and relationship aggression in college students. *Sex Roles, 38*, 743–756. https://doi.org/10.1023/A:1018868913615.

Saucier, D. A., O'Dea, C. J., & Strain, M. L. (2016). The bad, the good, the misunderstood: The social effects of racial humor. *Translational Issues in Psychological Science, 2*, 75–85. https://doi.org/10.1037/tps0000059.

Saucier, D. A., Strain, M. L., Miller, S. S., O'Dea, C. J., & Till, D. F. (2018). "What do you call a Black guy who flies a plane?" disparagement, confrontation, and failed subversion in the context of racial humor. *Humor, 31*, 105–128. https://doi.org/10.1515/humor-2017-0107.

Shurcliff, A. (1968). Judged humor, arousal, and the relief theory. *Journal of Personality and Social Psychology, 8*, 360–363. https://doi.org/10.1037/h0025493.

Strain, M. L., Martens, A. L., & Saucier, D. A. (2016). "Rape is the new black": Humor's potential for reinforcing and subverting rape culture. *Translational Issues in Psychological Science, 2*, 86–95. https://doi.org/10.1037/tps0000057.

Simmons, J. P., Nelson, L. D., & Simonsohn, Uri (2013). *Life after P-Hacking.* New Orleans, LA: Meeting of the Society for Personality and Social Psychology17–19. https://doi.org/10.2139/ssrn.2205186. Available at https://ssrn.com/abstract = 2205186.

Thomas, C. A., & Esses, V. M. (2004). Individual differences in reactions to sexist humor.

Group Processes & Intergroup Relations, 7, 89–100. https://doi.org/10.1177/1368430204039975.

Weaver, S. (2010). Developing a rhetorical analysis of racist humor: Examining anti-Black jokes on the internet. *Social Semiotics, 20*, 537–555. https://doi.org/10.1080/10350330.2010.513188.

Wicker, F. W., Barron, W. L., III, & Willis, A. C. (1980). Disparagement humor: Dispositions and resolutions. *Journal of Personality and Social Psychology, 39*, 701–709. https://doi.org/10.1037/0022-3514.39.4.701.

Wilson VanVoorhis, C. R., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorial in Quantitative Methods for Psychology, 3*, 43–50. https://doi.org/10.20982/tqmp.03.2.p043.

Wyer, R. S., & Collins, J. E. (1992). A theory of humor elicitation. *Psychological Review, 99*, 663–688. https://doi.org/10.1037/0033-295X.99.4.663.

Zillmann, D., & Bryant, J. (1994). Entertainment as media effect. In J. Bryant, & D. Zillmann (Eds.). *Media effects: Advances in theory and research* (pp. 437–461). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zillmann, D., & Cantor, J. R. (1976). A disposition theory of humor and mirth. In A. J. Chapman, & H. C. Foot (Eds.). *Humor and laughter: Theory, research, and application* (pp. 93–115). New York, NY: Routledge.