


RESEARCH ARTICLE

Rebuttals as deracializing strategies for African American and white candidates

Matthew Tokeshi 

Williams College, Schapiro Hall Room 238, 24 Hopkins Hall Drive, Williamstown, MA 01267, USA
Corresponding author. E-mail: mt13@williams.edu

(Received 24 January 2020; revised 13 April 2020; accepted 11 September 2020)

Abstract

The study of American racial politics has long focused on the conditions that activate racial animosity. A central line of research demonstrates that campaign messages that highlight negative stereotypes of African Americans can activate whites' racial attitudes. However, little is known about whether this activation can be overcome. I develop a theory of racial deactivation and test its predictions with two survey experiments. I find that explicitly criticizing the racial nature of an attack restores support for white candidates, but not African American candidates. However, African American and white candidates fare equally well using two rebuttal styles: a credible, non-racial justification of the attacked action or an explicit racial critique combined with the justification. The results have implications for how race affects campaigns, the susceptibility of the American public to racial cues, and campaign strategy.

Keywords: Race and elections; white voters; African American candidates

In 2006, Democrat Deval Patrick became the second elected African American governor since Reconstruction when he won the Massachusetts gubernatorial election over Republican Kerry Healey. Patrick was elected despite facing racially tinged campaign attacks that resembled the Willie Horton ads used by George Bush to paint former Massachusetts governor Michael Dukakis as soft on crime in the 1988 presidential election (Mendelberg 2001). One notable Healey ad began with a white woman walking alone in a dark parking garage and then cut to footage of Patrick describing a convicted rapist as “thoughtful” and “eloquent” (Cooper 2006). Patrick responded to this attack by pursuing a number of different strategies: offering a justification; touting his own tough-on-crime record; and arguing that Healey’s attacks were an attempt to divert attention away from her own record. Patrick’s polling lead shrank considerably after the racially tinged ads started airing. However, his rebuttals to Healey’s attacks coincided with a rebound in his poll numbers.

Twelve years after Patrick's successful gubernatorial campaign, Andrew Gillum launched a bid to become Florida's first African American governor in 2018. His opponent, Republican Ron DeSantis, warned voters not to "monkey this up" by electing Gillum, who he repeatedly attacked as corrupt and soft on crime (Jacobs 2018). President Trump acted as a DeSantis surrogate, once calling Gillum a "stone-cold thief" (Skoneki 2018). While Deval Patrick shied away from addressing race in his response to the parking garage ad, Gillum denounced DeSantis' "monkey" comments in openly racial terms, saying that DeSantis and Trump "no longer do whistle calls. They're now using full bull horns" (Jacobs 2018). Gillum went on to lose the election by four-tenths of a percentage point.

Gillum and Patrick represent different viewpoints in the debate over whether to directly confront racialized attacks in racial terms. Though different strategies have been used, little is known about whether rebuttals are effective at all, and if so, which ones and why. Filling this gap matters not only for understanding why candidates win or lose. It also clarifies whether political communication can play a role in reducing the significance of race as a consideration in candidate evaluation. Scholars of *racial priming* often paint a dim picture of the American public as being highly susceptible to manipulation by racial cues (Mendelberg 2001; Valentino *et al.* 2002). Concern regarding the susceptibility of the public to racial priming may be alleviated if white voters are receptive to rebuttals.

The standing prediction in the racial priming literature is that calling attention to the racial nature of an appeal—as Gillum did—is the most effective strategy. In this paper, I test this hypothesis for identical African American and white candidates and find mixed evidence: although it restores support for a white candidate, African American candidates do not enjoy the same benefit. I compare the effect of the racial rebuttal to the effect of a nonracial rebuttal strategy found to be effective by the literature on blame management strategies—a *justification* of the attacked action (McGraw 1991)—and find that the justification improves evaluations of both candidate types. I also test a third rebuttal that combines a nonracial justification of the attacked action with a racial denunciation and find that it has positive effects for candidates of either race. By demonstrating the effectiveness of justifications, I find that attacked candidates' electoral prospects are not irreparably damaged by negative racial attacks.

In addition to testing each rebuttal's effect on candidate evaluation, I also examine the role of racial attitudes in shaping candidate evaluation. Here, I find that both African American and white candidates are unable to change the minds of racially resentful white respondents when using the racial rebuttal. However, these same respondents evaluate the other two rebuttal types favorably regardless of the target's race. Taken as a whole, the results suggest that while race constrains African American candidates in important ways, they can nevertheless respond as effectively as white candidates to racial attacks by using the right rebuttal strategy.

Aside from its relevance to the racial priming literature, this study offers several other contributions to the study of public opinion and campaigns. First, unlike past studies on the effects of campaign messages (Johnston *et al.* 2004; Gerber *et al.* 2011), I consider not only the effects of messages, but also how those effects change when pitted against a countermessage. This follows work by Chong and

Druckman (2007) and Sniderman and Theriault (2004), who study elite communication in this more realistic iterative framework. Second, this study not only adopts but also advances Chong and Druckman's (2007) counterframing paradigm by addressing how respondents evaluate counterframes that raise more than one consideration, which is commonly found in real campaigns but rarely tested in a controlled experimental setting (Chong and Druckman 2011). Third, although many scholars have examined the effects of negative campaigning (Ansolabehere and Iyengar 1995; Geer 2006), few have explored how the findings from that research are affected by a candidate's race (although see Krupnikov and Piston 2015). Finally, I compare identical messages offered by African American and white candidates, an approach that builds on rare studies of the effects of source cues (Kuklinski and Hurley 1994; Nelson *et al.* 2007). While informative, these studies do not consider how source cues operate in an iterative electoral context, as is done here.

1. Previous literature

Implicit racial appeals—ones that avoid direct reference to race, yet still communicate racial messages by alluding to racial stereotypes or using racial imagery—remain common in U.S. politics.¹ A systematic examination of campaign ads in U.S. House and Senate races conducted by McIlwain and Caliendo (2011) finds that racial stereotypes and imagery were common in attack advertising aimed at non-white candidates in the 1990s and 2000s. Research demonstrates that such cues lead to the activation of racial attitudes among white voters, making racial attitudes a central determinant of citizens' opposition to the representation of African Americans and their interests (Kinder and Sanders 1996; Mendelberg 2001; Valentino *et al.* 2002; Hurwitz and Peffley 2005; White 2007; Tesler and Sears 2010; Kinder and Dale-Riddle 2012; Tesler 2012).

Although many studies have documented the effects of implicit appeals, few have examined the effects of rebuttals to them. Early studies found that white Democrats in the 1970s and 1980s did not directly rebut implicit appeals made against them, instead choosing to ignore them or change the subject (Edsall and Edsall 1991; Glaser 1996). Studies of African American candidates in the 1980s also found that explicit discussion of race, including directly challenging implicit appeals as racist, was not a popular strategy (see Perry 1991 on "deracialization").

The first test of the effect of rebuttals to implicit appeals was conducted by Mendelberg (2001) using survey data from the 1988 presidential campaign. Using the date of interview from the 1988 American National Election Study (ANES) as a measure of exposure to the Horton ads, Mendelberg finds that racial attitudes had the smallest effect on candidate evaluations during the last phase of the campaign in which the Horton ads were explicitly called anti-black messages. Thus, Mendelberg argues that calling attention to the racial nature of the appeal is a successful rebuttal strategy. However, this finding is open to alternative explanations. For example, since message exposure is measured indirectly (using date of interview as a proxy), it is unclear whether the message itself or some other factor that occurred at the same time caused candidate evaluations to change. An experimental test is needed to determine the causal relationship between message and candidate evaluation. Also,

Mendelberg's case study examined the response of a white presidential candidate's campaign. Black candidates' responses have never been tested.

2. Rebuttals to racial appeals

Research in political communication provides insight into how rebuttals might play an important role in combatting racial attacks. The pioneering work on “counterframing” by Chong and Druckman (2007) and Sniderman and Theriault (2004) critiques earlier work on framing, which showed that the way in which political arguments are “framed” can have dramatic effects on the types of considerations (such as group attachments, values, and interests) that citizens draw upon when making political evaluations (Kinder and Sanders 1996; Nelson and Kinder 1996; Nelson *et al.* 1997). Chong and Druckman (2007) criticize the early framing research for ignoring the reality of democratic competition among elites. They argue that in the real world, opposing elites offer competing frames to the public in an effort to persuade citizens to evaluate a political matter using their preferred set of criteria.

In their experiments, they find that framing effects dissipate when pitted against a counterframe that emphasizes a different consideration, assuming that the counterframe is a compelling argument. This occurs because the counterframe promotes deliberate evaluation of the original frame. Though race and racial attacks are not discussed in their work, the relevance of their findings for this study's examination of rebuttals is clear: while racial attacks are a type of frame that is meant to activate racial considerations, rebuttals are a type of counterframe that can potentially raise an alternative set of considerations, thus changing how citizens evaluate candidates and the considerations they draw upon when doing so.

What remains unclear in the Chong and Druckman paradigm is what constitutes a compelling (or “strong”, in their terminology) frame. I advance their work by testing three counterframes to attacks that activate anti-black sentiment offered by identical black and white candidates. In the next section, I explain key features of each rebuttal and highlight other areas of political communication research that offer clues for how they might work. In the section after that, I consider how the race of the messenger may influence each rebuttal's effect.

3. Rebuttal strategies

The first strategy tested is the *racial* rebuttal. This was the one that Mendelberg (2001) found effective—calling attention to the racial nature of the attack. Barack Obama used this strategy during the 2008 campaign in response to a John McCain ad that interspersed Obama images with irrelevant images of the blonde female celebrities Paris Hilton and Britney Spears. Obama replied to the ad by saying that it was an attempt to remind voters that he “doesn't look like all those other presidents on the dollar bills” (Mooney 2008).

The standing prediction in the racial priming literature is that calling attention to the racial nature of the appeal is the most effective rebuttal strategy. As mentioned earlier, Mendelberg (2001) shows that racial attitudes had a significantly weaker effect on candidate evaluations during the last phase of the 1988 presidential campaign in

which the Horton ads were explicitly called anti-black messages by leading Democrats, a charge that was widely disseminated by the media. The reason for this weakening racial effect, according to this account, is that the rebuttal “uncovered” the racial content of the Horton ad, which led whites to reject the message since it violated a widely held norm against openly racist derogation. The rebuttal should work even among racially resentful whites since supposedly even they, too, have internalized this norm. Thus, a message that would have worked as long it remained implicitly racial will no longer do so once it is uncovered as an explicitly negative reference to African Americans. From this perspective, *racial* works precisely because it is highlighting the racial character of the implicit message.

On the other hand, there are reasons to believe that *racial* may be ineffective or even counterproductive. One reason for skepticism comes from scholars in the “deracialization” tradition, who argue that African American candidates who avoid explicit discussion of race stand the greatest chance of success (Hamilton 1977; Citrin *et al.* 1990; Sonenshein 1990; Perry 1991; Jeffries and Jones 2006; Frederick and Jeffries 2009). A deracialized style is considered beneficial because it attempts to minimize the salience of racial considerations as much as possible. Thus, directly challenging an implicit appeal as racist makes little sense because it reinforces considerations that the candidate should be trying to suppress. From this perspective, *racial* will not work because either whites may not find charges of racism to be credible or the norm of racial equality no longer acts as a constraint on their response, a hypothesis that is supported by recent research (Valentino *et al.* 2017).

Although *racial* has received the most attention in the racial priming literature, other rebuttal types are possible and suggest alternative mechanisms for explaining whether citizens revise their evaluations of public officials who are accused of wrongdoing. The second strategy tested in this experiment—*justify*—involves explaining the attacked action by tying together relevant facts into a reason-based narrative. For example, former Arkansas governor Mike Huckabee employed this strategy in 2009 when he was criticized for pardoning Maurice Clemmons, an Arkansas convict who murdered four Washington state police officers after Huckabee pardoned him. Huckabee justified the pardon by saying that he received a recommendation to pardon Clemmons from the trial judge in the case and the unanimous vote of the parole board (Sterling 2009).

Evidence consistent with the notion that *justify* can be persuasive is found in the political accounts literature begun by McGraw (1991), who demonstrated the effective use of politicians’ reason- and fact-based justifications for their controversial actions. According to McGraw (1991), a justification does not deny responsibility for the action. Instead, “a justification focuses on the outcome and claims that contrary to accusations, the consequences are not so undesirable and therefore that less or no blame is warranted (1136).” Thus, a successful justification should lead to a two-step inferential process. First, if the citizen is satisfied with the justification, it should lead to positive reevaluations of the policy in light of the public official reframing its consequences in a more positive manner. Second, as a result of this reframing, the citizen may credit the public official for the policy action. While *racial* suggests that attacked public officials must “unmask” the racial intent of an attack, *justify* focuses on the role of relevant and credible information in counteracting an attack’s harmful effects.

A third possible rebuttal strategy involves a more nuanced type of racial response—one that combines racial and compelling nonracial themes. In this study, this strategy is conceptualized as *justify + racial*, a hybrid response that combines a justification with a racial reference. One example of this response was Barack Obama’s “A More Perfect Union” speech from March 2008 in which he addressed the controversial videos that emerged of his former pastor Jeremiah Wright denouncing the United States. In the speech, Obama tried to *justify* the origins of his relationship with Wright and how his former pastor, like many people, has strengths as well as shortcomings. Akin to *racial*, Obama acknowledged the existence of white racism in American society in a way that many observers considered unusual for a presidential candidate at the time (Kinder and Dale-Riddle 2012, 45).

By disaggregating racial messages into subtypes based on their content, this study builds on the work of Stout (2015), who distinguishes between *negative* racial appeals—ones that “are generally devoid of substantive policy proposals and usually take the form of an attack on a political opponent, the media, or a supposed ally” (11–12)—and *positive* racial appeals—ones that “either advance black policy interests or highlight the candidate’s connection to the black community without attacking outside political players or institutions” (8). For Stout (2015), tone is the key factor in determining whether whites accept racial rhetoric. He theorized that negative racial messages backfire among white voters because many view charges of racism as unnecessarily harsh. However, they may be open to racial messages that do not focus exclusively on charges of racism. Perhaps relevant and credible information may soften the tone of racial rebuttals, thereby making them more persuasive to white respondents.²

In addition to providing a test of a more nuanced style of racial response, *justify + racial* also allows for a test of whether the effects of *racial* and *justify* are conditional on the other. In other words, it allows for an assessment of whether *racial*’s effect is conditional on it being packaged with a compelling non-racial consideration such as *justify*, in line with Stout’s (2015) finding that racial messages that focus only on highlighting racism are less effective than other types of racial messages. It also allows for a test of whether *justify*’s effect is conditional on it being strictly non-racial.

4. Candidate race

In addition to the content of the rebuttals themselves, the race of the messenger is likely to influence how rebuttals are received (Kuklinski and Hurley 1994). Tests of identical messages offered by African American and white politicians have found a number of instances in which messages offered by African American politicians are received less positively. For example, Nelson *et al.*, (2007) found that African American elites’ charges of racism are perceived as less credible than the same charges made by white elites. The implication of this finding for the current study is that since many observers may not view implicit appeals as being racial in nature, the effectiveness of the racial rebuttal hinges on whether the accusation of racial intent is perceived as credible. If African American claims of racism are perceived as less credible, then the racial rebuttal is also less likely to be effective when used by an African American candidate.

Other research suggests that the rhetorical penalty suffered by African American candidates extends beyond racial messages. Studies of negative advertising (Krupnikov and Piston 2015) and ambiguous rhetoric (Piston *et al.* 2018) find that white candidates enjoy greater payoffs from using these nonracial communication strategies. Thus, we may observe racial differences in the effect of a nonracial rebuttal such as a justification.

Aside from rebuttals' main effects, another important outcome to examine is whether rebuttals can diminish the effects of racial attitudes on candidate judgments. Evidence consistent with the notion that rebuttals can deracialize is found in the literature on counterframing cited earlier (Sniderman and Theriault 2004; Chong and Druckman 2007). Applying their insight about counterframes raising alternative considerations to the current study, rebuttals that emphasize compelling nonracial considerations (such as a justification) may succeed in shifting focus away from the racial considerations that were primed by the attack.

Optimism regarding a rebuttal's potential to deracialize an attacked candidate is likely to apply more to white candidates than to African American candidates, however. Previous research on white candidates finds that the extent to which they are evaluated through a racial lens depends on the nature of the campaign communications environment (Kinder and Sanders 1996; Mendelberg 2001; Valentino *et al.* 2002). Research on African American candidates, however, suggests that by virtue of their race alone, they make racial considerations "chronically accessible" in political decision-making (Tesler and Sears 2010). Hajnal (2007, 16), in a study of white attitudes toward African American mayoral candidates, finds that "black challengers can and usually do try to counter the uncertainty surrounding their candidacies by running "deracialized" or pro-white campaigns, but white voters tend to ignore these campaign statements, which they perceive as having little credibility." Thus, rhetorical strategies aimed at shifting attention away from racial considerations (such as a non-racial justification) are unlikely to have their intended effect for an African American candidate.

5. Research design

To test the effects of rebuttals to implicit racial appeals, I conducted two internet survey experiments. The first experiment, conducted in March and July 2013, used a convenience sample of 265 white respondents from Amazon.com's Mechanical Turk (MTurk). While MTurk respondents are more diverse than student and convenience samples that are often used in political science experiments (Berinsky *et al.* 2012), they are not a nationally representative sample of American adults. To help compensate for this limitation, a second experiment was conducted in February 2014 using a sample of 380 white respondents provided by Survey Sampling International (SSI) that was drawn to match the U.S. white adult population on age, sex, and region. Results from the two experiments were substantively similar, so the two samples are combined in the following analyses.³ Combining the samples yields a total of 645 white respondents.⁴

The designs of the two experiments were identical. Both simulated an attack-rebuttal episode in a fictitious U.S. Senate campaign using a within-subjects design.

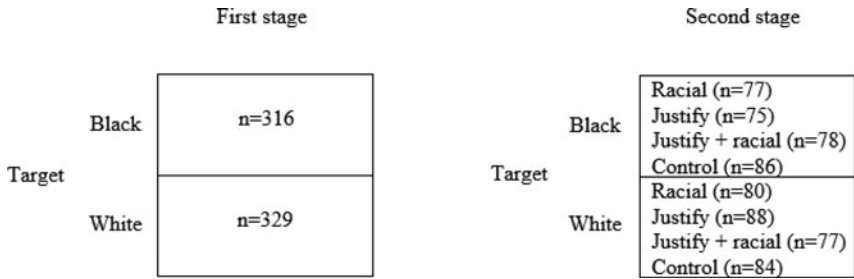


Figure 1. Experimental design (Note: While the target’s race is either black or white, the criminal’s race is black in all conditions).

Respondents were randomized into eight conditions: 1 (attack) × 4 (rebuttal types: *racial*, *justify*, *justify + racial*, or no rebuttal) × 2 (target race: black or white).

Figure 1 shows the basic design of the experiments. They proceeded in two stages. In the first stage, respondents read a fictitious news story about an attack ad in a U.S. Senate race. The story explains that the target of the ad was attacked by his opponent for pardoning a former U.S. Representative (the “criminal”) who was convicted of eight felonies, including assault of a police officer that left the officer hospitalized due to head injuries. After reading the news story, respondents provided feeling thermometer assessments (ranging from 0 = negative to 100 = positive) of the target and the attacker and answered a battery of demographic questions (but none about race so that responses to remaining questions are not primed by race). In the second stage, respondents were randomly assigned to one of the four rebuttal conditions (*racial*, *justify*, *justify + racial*, or control) and offered a second round of feeling thermometer evaluations for both candidates. To minimize the possibility of priming racial attitudes before reading the target’s rebuttal, respondents answered the four-question racial resentment battery (Kinder and Sanders 1996) after the second round of candidate evaluations.⁵

The pardon of a convicted ex-politician was selected as the theme of the ad described in the first stage news story because linking candidates to black criminality has been a theme in campaign ads such as the Willie Horton ad, the anti-Deval Patrick parking garage ad, and an ad from 2008 that tied Obama to Kwame Kilpatrick, the African American former mayor of Detroit who pleaded guilty to two counts of obstruction of justice and later resigned. The scenario examined in this study is intended to be similar to the Kilpatrick attack, as the ex-Representative’s crimes are loosely modeled on Kilpatrick’s. A side-by-side photo of the target and the criminal is embedded in the news article (see Figure 2). The target is shown wearing a suit while the criminal is shown in a mug shot. The race of the target was manipulated to be either white or black in order to test whether race affects rebuttal effectiveness. The criminal’s race is black in all conditions.

The white and black versions of the target must be as close to identical as possible so that differences in ratings between them can be attributed to their skin color and not some other aspect of their appearance, such as their attractiveness. In order to do

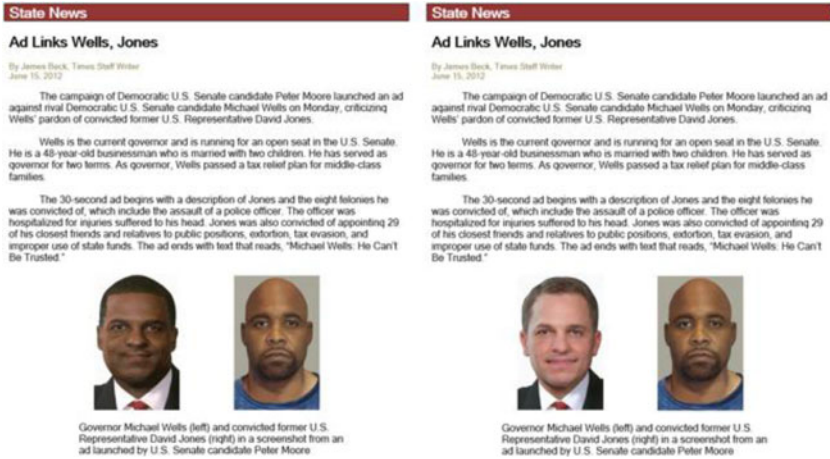


Figure 2. Two versions of the attack news story, black target (left) and white target (right). The text of the news story is shown in Figure 1 of the Appendix.

this, a morphing procedure used in earlier political science research was used (Bailenson *et al.* 2008; Weaver 2012). This procedure yields a white face and a black face that are 60% identical, which controls for many extraneous (nonracial) sources of variation while still allowing realistic variation in skin tone and facial features.⁶

In order to isolate the effect of race and neutralize interactions with party identification, I assign respondents to a primary contest between the target and attacker that matches the respondent’s partisanship.⁷ In other words, Democratic respondents are assigned to a Democratic primary; Republicans are assigned to a Republican primary; and Independent respondents are randomly assigned to either a Democratic or Republican primary.

Turning to the second stage of the experiment, Table 1 highlights the key features of the four rebuttal conditions.⁸ All conditions (except the control condition) contain the same first paragraph, explaining that a U.S. Senate candidate attacked his opponent for pardoning a convicted former U.S. Representative. The second and third paragraphs vary for each of the rebuttal conditions:

- In *racial*, the target says, “This ad is an attempt to stir up racial fears...Charges like this breed division in our country and our state. They divide us—*race against race*—so we blame each other instead of work together” (emphasis added). This statement is meant to resemble the one offered by then-Governor Bill Clinton’s speech announcing his run for the presidency in 1991 when he said, “For twelve years, Republicans have tried to divide us—*race against race*—so we get mad at each other and not at them” (Mendelberg 2001, 104).
- The target in *justify* offers a justification for his action. He explains that the trial judge in the case recommended commuting the sentence and that the convicted

Table 1. Description of rebuttal stories

Rebuttal	<i>Racial</i>	<i>Justify</i>	<i>Justify + racial</i>	Control
Headline	Wells Claims Ad Had Racial Intent	Wells Claims Ad is Unfair Attack	Wells Claims Ad is Unfair, Racial Attack	YouTube Takes On Television
First Paragraph	Wells fired back against ad...	Identical across conditions	Identical across conditions	Information about a non-political news story
Second Paragraph	"...ad is an attempt to stir up racial fears..."	"...ad is a distortion of the truth..."	"...ad is a distortion of the truth..."	Information about a non-political news story
Third Paragraph	"Charges like this... divide us—race against race..."	Wells explains the reason for pardon: unanimous vote of bipartisan review board	Wells explains the reason for pardon, plus: "Charges like this... divide us—race against race..."	Information about a non-political news story
Photos	Candidate (black or white) and criminal (black)	Identical across conditions	Identical across conditions	Photos that go with a non-political news story

Representative received a 5-0 vote recommending parole from the state's bipartisan parole board. He concludes by saying that he does not condone the Representative's actions but relied on the good judgment of the trial judge and parole board. This reasoning was similar to the real-life justification offered by Mike Huckabee described earlier.

- *Justify + racial* combines the justification offered in *justify* with *racial*'s claim that charges like this divide us "race against race."
- Finally, respondents read a non-political news article in the control condition that is similar in length and formatting to the other rebuttal conditions. After reading the article, respondents evaluate the target and the attacker they read about in the first stage news story. Thus, all respondents evaluate the fictional candidates twice. The control condition is needed in order to account for the possibility that the mere act of repeated measurement may influence treatment effects due to anchoring, learning, fatigue, or some other time-related phenomenon. In other words, if a change in candidate evaluations is observed in one of the rebuttal conditions but no change is observed in the control condition, we can be confident that the change in the rebuttal condition is not an artifact of repeated measurement.

6. Results

Following studies in the counterframing and political accounts literatures, the analysis focuses on how rebuttals affect candidate evaluation. In addition to these main effects, I also address studies in the racial priming literature by examining the impact of racial resentment in shaping candidate evaluation. This allows for a test of whether rebuttals can help candidates deracialize.

The dependent variable used to test the main effect of a rebuttal is the change between the first round (post-attack) and second-round (post-rebuttal) of candidate feeling thermometer ratings. It is measured by (1) calculating the target-attacker feeling thermometer difference after the attack; (2) calculating the same difference after the rebuttal; and (3) calculating the difference between the second quantity and the first quantity and dividing by 100 to map on the -2 to 2 scale.⁹ This score is the measure of the improvement in the evaluation of the target caused by the rebuttal with positive values indicating that the rebuttal helped the target, zero indicating that the rebuttal had no effect, and negative values indicating that the rebuttal backfired. The target-minus-attacker differences are used instead of simply looking at the target's ratings in order to account for the possibility that the attacker might suffer a penalty for issuing a negative attack (Lau and Rovner 2009).

Main effects are tested in three ways. First, is the effect positive and significantly different from the control group? Second, does the race of the target matter for the effect of a given rebuttal? After I conduct tests 1 and 2 for each rebuttal, I conduct test 3: which rebuttal is most effective for each target? These tests jointly cover all possible paired comparisons in a 4×2 design.

Figure 3 shows each rebuttal tested against a null hypothesis of no change relative to the control group (test 1) and all comparisons between pairs (tests 2 and 3). Each

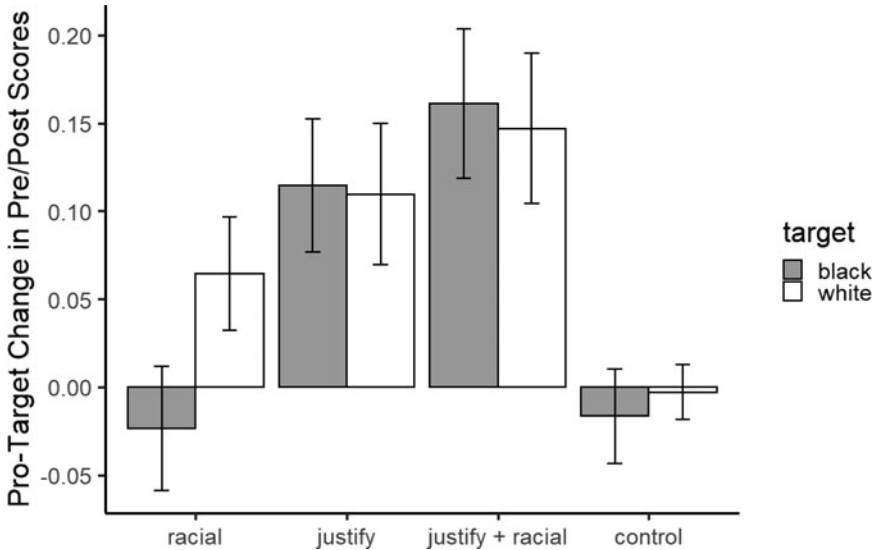


Figure 3. The impact of rebuttals on target-attacker evaluations. Positive values indicate that the rebuttal helped the target. Error bars represent 84% confidence intervals, which is consistent with a two-tailed test of overlap at $p = 0.05$. In other words, values are significantly different where error bars do not overlap. The “rebuttal effect” is the average change in target-attacker rating from post-attack to post-rebuttal for each rebuttal.

bar represents the rebuttal effect for each target-rebuttal pair with its 84% confidence interval, which is consistent with a two-tailed test of overlap at $p = 0.05$. In other words, values are significantly different where confidence intervals do not overlap. To aid in interpreting the magnitude of the rebuttal effects, the attack resulted in a target rating that was 0.14 lower than the attacker’s rating on average.¹⁰ Therefore, rebuttal effects can be interpreted as the amount of the 0.14 hit that the target gets back by using the rebuttal.

I start with *racial*. Looking at the average magnitudes compared to the control group (test 1), we see that it improves ratings of a white target by an average of 0.07 ($p < 0.01$), but the rebuttal’s effect when used by a black target is statistically indistinguishable from zero ($p = 0.82$).¹¹ A direct test of the racial disadvantage hypothesis (test 2) indicates that support for the white target is 0.08 higher on average than for a black target using the same rebuttal ($p = 0.01$). Overall, the results provide mixed support for the racial priming literature’s endorsement of *racial*. While it improves the evaluation of a white target, a black target does not enjoy the same benefit. This is consistent with previous research on the effects of racial rhetoric for black versus white elites (Nelson *et al.* 2007).¹²

Turning to *justify*, Figure 3 shows that this rebuttal improves evaluations of black and white targets alike (test 1). *Justify* improves evaluations by an estimated 0.13 for a black target and 0.11 for a white target compared to each target’s appropriate control group ($p < 0.01$ for both tests). The difference in *justify*’s effectiveness by candidate race is not statistically distinguishable from zero (test 2) ($p = 0.90$). In sum, *justify*

is effective for a target of either race. This is also the case for *justify + racial* as it boosts evaluations of a black target by 0.18 and a white target by 0.15 compared to the relevant control group (test 1) ($p < 0.01$ for both tests). The difference between black and white targets is also statistically indistinguishable from zero (test 2) ($p = 0.74$) even though the rebuttal includes a mention of race.

Next, I turn to test 3: which rebuttal works best for each target? Starting with the black target, we learned from test 2 that *justify* and *justify + racial* were successful relative to the control while *racial* was not. Comparing *justify* to *justify + racial*, there is no statistically significant difference ($p = 0.25$), while both are more effective than *racial* ($p < 0.01$ for both tests). Based on test 3, it appears that the effective set of strategies for a black target are *justify* and *justify + racial*. At the same time, a black target should avoid *racial*, though it is worth noting that *racial* does not cause a backlash effect compared to the control group. The rebuttal simply leaves a black target no better or worse off than not responding at all.

For the white target, test 2 revealed that all three rebuttals were successful relative to the control group. Turning to test 3, *justify + racial* was more effective than *racial* ($p = 0.03$), but indistinguishable from *justify* ($p = 0.22$). *Justify* and *racial* were indistinguishable from each other ($p = 0.22$). Thus, as with the black target, the set of two strategies that cannot be dominated by any other are *justify* and *justify + racial*. The difference for a white target is that *racial* is also effective, though less so than *justify + racial*.¹³

To better understand the main effects of each rebuttal, I examine whether rebuttal effects persist for respondents with high levels of racial resentment. The racially resentful provide a hard test for rebuttals. If rebuttals have a positive effect on the racially resentful, this would suggest that rebuttals can deactivate racial attitudes—a normatively attractive result given the inherently manipulative aspect of racial priming. Figure 4 shows the average rebuttal effect among high racial resentment respondents only (those in the top third of the sample, all of whom scored at or above 0.6875 on the 0–1 racial resentment scale ($n = 245$)).¹⁴

Starting with *racial*, we see a key difference from the overall findings. Compared to the control group, a white target using *racial* no longer enjoys statistically significant improvement in his evaluation ($p = 0.48$). As in the overall sample, a black target using *racial* is not significantly better off than the black target in the control group ($p = 0.71$). Comparing the white target to the black target directly, support for the white target is 0.11 higher on average, a marginally significant difference ($p = 0.05$). In sum, we see that racially resentful respondents do not respond positively to *racial* for a target of either race, but they especially do not like it when used by a black target.

Black targets, however, fare better among the racially resentful by using *justify*. A black target using *justify* enjoys 0.19 more support on average than a black target in the control group ($p < 0.01$). A white target using *justify* averaged 0.06 more support on average than a white target in the control group, but the difference is not statistically significant ($p = 0.26$). The comparison between black and white targets using *justify* yielded no significant difference ($p = 0.33$). Racially resentful respondents also appear to respond positively to *justify + racial*. Mirroring the overall results, *justify + racial* is more effective for both black and white targets compared to their respective control groups (black difference = 0.18, $p < 0.01$; white difference = 0.15,

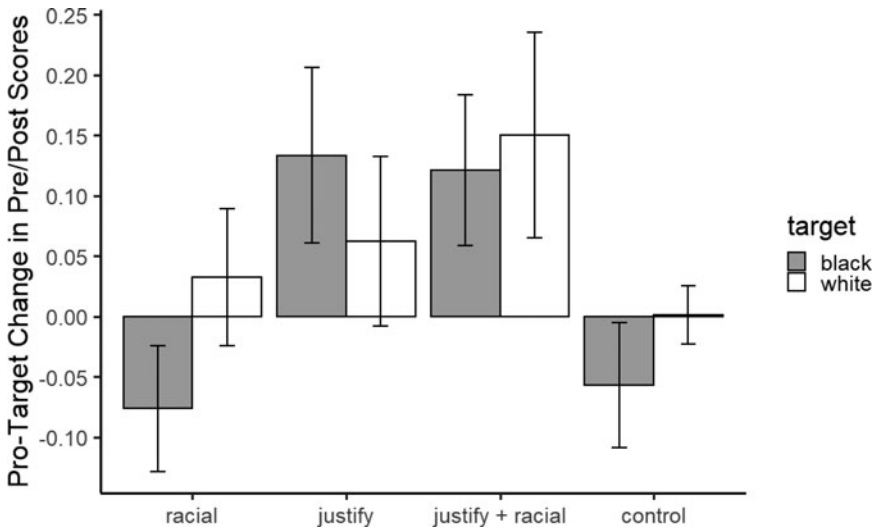


Figure 4. Rebuttal effects among high racial resentment respondents only (top third of the sample, corresponding with racial resentment ≥ 0.6875 on a 0–1 scale). Positive values indicate that the rebuttal helped the target. Error bars represent 84% confidence intervals, which is consistent with a two-tailed test of overlap at $p = 0.05$. In other words, values are significantly different where error bars do not overlap.

$p = 0.02$). The difference between black and white targets is not statistically significant ($p = 0.70$).

Overall, the results among the racially resentful are surprisingly similar to the results for the overall sample. The key difference is that *racial* no longer appears to benefit a white target. This suggests that a rebuttal that criticizes a racialized attack in explicitly racial terms does not necessarily lead racially resentful whites to reevaluate the racialized attack, contrary to Mendelberg's (2001) theory of internalized racial norms. On the other hand, racially resentful whites appear to respond more positively to rebuttals that offer credible and relevant information even when that information is paired with explicitly racial criticism. Unlike *racial*, this holds for both black and white targets. In sum, results from this analysis suggest that information and tone are more important influences on candidate evaluation than highlighting norm violation among whites with high levels of racial resentment.

Finally, I examine rebuttal effects among the racially sympathetic, defined as respondents in the bottom third ($RR \leq 0.5$) of the racial resentment distribution ($n = 273$).¹⁵ The results, presented in Figure 5, resemble the results for the overall sample shown in Figure 3.¹⁶ The white target using *racial* enjoys a marginally significant effect of 0.07 relative to the white target in the control group ($p = 0.06$). In light of *racial*'s lack of an effect for white targets among high racial resentment respondents, it appears that its effect among the overall sample is driven largely by racially sympathetic respondents. Meanwhile, *justify* and *justify + racial* are more effective for both black and white targets compared to their respective control group ($p = 0.07$ for black *justify* versus black control; $p < 0.01$ for all other comparisons). Overall, the

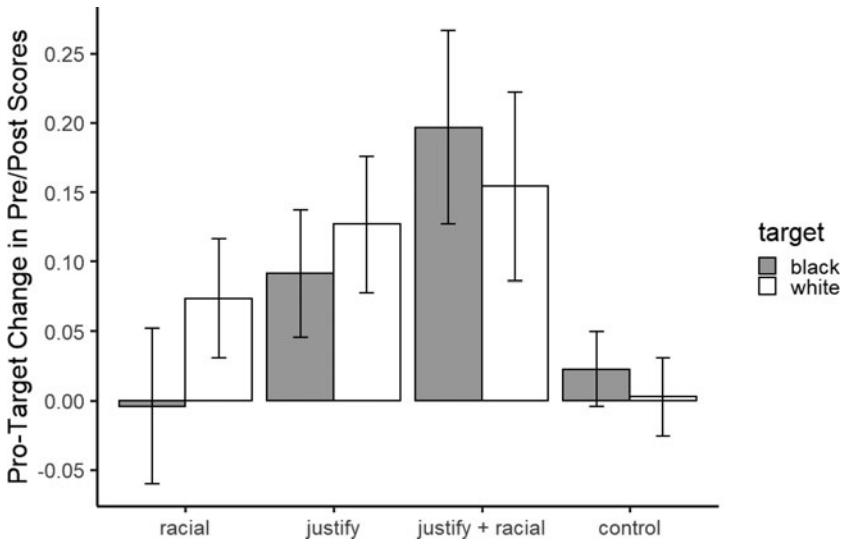


Figure 5. Rebuttal effects among low racial resentment respondents only (bottom third of the sample, corresponding with racial resentment ≤ 0.5 on a 0–1 scale). Positive values indicate that the rebuttal helped the target. Error bars represent 84% confidence intervals, which is consistent with a two-tailed test of overlap at $p = 0.05$. In other words, values are significantly different where error bars do not overlap.

analysis by racial resentment level suggests that while no racial resentment subgroup approves of a black target using *racial*, a white target offering the same rebuttal can mobilize support among the racially sympathetic. The other rebuttals work well across the racial resentment spectrum and without differences by target race.

7. Discussion

Despite numerous examples of racial attacks in recent campaigns, little is known about whether rebuttals are effective and, if so, which ones and why. In this paper, I define and systematically test three rebuttals used by African American and white candidates. Answering this question is important not only for understanding electoral outcomes but also for understanding the limits of racial campaigns and the public’s capacity to respond to credible communication. I find that calling attention to the racial nature of the attack benefits a white candidate, but not an African American candidate, contrary to the standing prediction in the racial priming literature. However, two rebuttal types proved effective for white and African American candidates alike—(1) a non-racial justification of the attacked action and (2) a modified version of racial defense that includes a non-racial justification.

The key intellectual contribution of the current study pertains to how African American candidates experience and combat racial appeals, a departure from typical research on racial appeals, which focuses on the effects of racial appeals on white candidates (e.g., Mendelberg 2001). The findings have negative and positive implications for African American candidates facing racial appeals. On the negative side, there is

clear evidence of continuing racial discrimination. The most obvious place to see this is in the responses to the *racial* rebuttal. In spite of this disadvantage, however, the findings also point to the more optimistic conclusion that African American candidates can still neutralize the damaging effects of a racial attack. The success of justifications is in line with studies that portray white voters, even racially resentful ones, as receptive to reasonable information (Hajnal 2007; Goldman and Mutz 2014). The success of the hybrid *justify + racial* also suggests that adopting a strictly deracialized style in the face of racial attacks may be unnecessary. To be sure, black targets who only highlight the racial nature of the attack fare poorly. But black targets can succeed with nuanced rebuttals that highlight compelling racial and non-racial themes. In short, the packaging of racial defense matters.

The results of this study also speak to an important theoretical question in the study of American racial politics. A key debate that has emerged during the Obama and Trump presidencies is whether the norm against explicit racial derogation still acts as a constraint on white Americans. Mendelberg (2001) provided evidence that even whites with high levels of racial resentment reject explicitly racialized rhetoric, an argument that has been challenged by recent research suggesting that many whites no longer consider explicitly hostile racial rhetoric to be inappropriate (Hutchings *et al.* 2010; Valentino *et al.* 2017). The findings in this study support the view of the newer research. While *racial* improved evaluations of the white target among the overall sample, the analysis by racial resentment level shows that only the racially sympathetic respond positively. Racially resentful whites were not moved by a rebuttal that made visible a message that likely would have violated consensus norms in earlier decades.

One limitation of this study is that it focuses only on the opinions of white respondents. The response of African American voters would factor into any politician's thinking about the strategic use of rebuttals, and therefore, future research should take this into account. Another limitation is that the data was collected before Donald Trump became president, and so it is worth considering how these findings would hold up in the Trump era. Given recent findings that suggest white liberals have grown more racially sympathetic since the 2016 election (Pew 2017; Englehardt 2019; Jardina 2019), it is possible that African American candidates charging racism may have a more receptive audience than the one found in this study. A third limitation is that other realistic rebuttal strategies were not tested. For example, a candidate can choose to ignore the attack and pivot to a topic more favorable to him such as policy positions or accomplishments. This is the strategy employed by Barack Obama for most of the 2008 presidential campaign. Obama usually avoided directly addressing racially tinged attacks or media speculation about the role of race in the campaign, instead choosing to focus on his economic message (Tesler and Sears 2010, 54–57). Future research should examine whether African American candidates can successfully fend off racial attacks by changing the subject in this manner.

It is likely that racial attacks like the ones Andrew Gillum faced in 2018 will remain a significant part of campaigns for the foreseeable future, so much work on rebuttals remains to be done. This study is a first step in addressing the important questions of which strategies work and for whom. Aside from these practical considerations, this

study also has normative implications for the quality of public opinion in U.S. politics. Racial politics scholars often portray the American public as being easily manipulated by racial cues. This study builds on the existing literature by showing that implicit racial attacks do not necessarily determine the role of race in campaigns. Instead, African American and white candidates can counter the negative effects of these attacks depending on how they respond.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/rep.2020.34>

Acknowledgements. I thank Zack Chauvin for research assistance. For feedback and suggestions on previous versions, I thank Tali Mendelberg, Martin Gilens, Victoria Shineman, and the anonymous reviewers.

Notes

1 The literature on racial appeals draws a distinction between ones that attempt to activate racial group identity and mobilize support among racial minority constituencies (White 2007; McIlwain and Caliendo 2011; Stout 2015) and ones that aim to draw on anti-black sentiment for the purpose of influencing white voters (Mendelberg 2001; Hutchings *et al.* 2010). This paper focuses on the second (anti-black) type of racial appeal.

2 While racial priming studies focus primarily on how messages influence the association between racial resentment and candidate evaluation, Stout (2015) examines how messages influence candidate evaluations. This is in line with the counterframing and political accounts literatures discussed earlier. The analysis presented later examines how messages influence candidate evaluations (the main effects) and racial resentment's role in shaping those evaluations.

3 For a full table of the rebuttal effects in each experiment, see Table 1 in the Appendix.

4 Although the sample skews younger than a nationally representative sample would due to the inclusion of the MTurk respondents, there is still notable variance across key socioeconomic variables and political attitudes. Women were 46% of the sample. Democrats made up 40% of the sample, while Republicans accounted for 34%. The median age was 38 years old, while the median annual household income category was \$50,000–\$74,999. For more on sample demographics, see Table 2 in the Appendix. Table 3 in the Appendix shows the results of a balance test, which confirms that no significant differences exist between treatment groups for age, income, education, and partisanship. The percent male was significantly lower in black *racial* compared to black *justify* and white *justify*. However, rebuttal effects did not significantly differ by gender for any of the rebuttal conditions, as shown in Appendix Table 4.

5 To address concerns that measuring racial resentment after the second round of candidate evaluations opens the possibility of post-treatment bias, the average racial resentment score of respondents assigned to a black target was not significantly higher than that of respondents assigned to a white target ($p = 0.34$, two-tailed test).

6 For more details on how the faces were created, see Tokeshi and Mendelberg (2015).

7 Respondents' party identification is asked before they receive any treatment so that they can be assigned to the appropriate primary.

8 See Figures 2–5 of the Appendix for the full text of each rebuttal.

9 For example, imagine a respondent who gave post-attack ratings of 30 to the target and 50 to the attacker and post-rebuttal ratings of 40 to the target and 35 to the attacker. The change in candidate thermometer ratings would be $(40 - 35, \text{ the post-rebuttal difference}) - (30 - 50, \text{ the post-attack difference}) = 25$, which divided by 100 equals 0.25, a strong endorsement of the target's rebuttal. Although the theoretical range of this variable is -2 to 2 , the actual range for the sample was -1 to 1.2 , with most of the data clustering between -0.01 and 0.15 (the interquartile range).

10 The attack results in an average target rating that is 0.13 lower than the attacker's rating for black targets compared to 0.15 lower for white targets. This difference is statistically indistinguishable from zero ($p = 0.62$).

11 All tests reported in this paper are two-tailed.

12 Figures 6, 7, and 8 in the Appendix show the results for Democrats, Republicans, and Independents, respectively. With respect to *racial*, interesting partisan differences emerge. Among Democratic respondents (Figure 6), black targets do not fare worse than white targets when using *racial* ($p = 0.24$). While black targets using *racial* do not enjoy a boost relative to the black control group ($p = 0.22$), white targets using *racial* are better off than white targets in the control group ($p < 0.01$). Among Republican respondents (Figure 7), black targets using *racial* fare worse than white targets using the same rebuttal ($p = 0.03$), mirroring the racial penalty observed among the overall sample. Black targets using *racial* are also marginally worse off than black targets in the control group ($p = 0.07$). Though white targets are better off than black targets using *racial*, they are not better off compared to the white control group ($p = 0.38$). In sum, the overall null effect of a black target using *racial* appears to be an average of a mildly positive effect among Democrats and a strongly negative effect among Republicans, while the overall positive effect of a white target using *racial* is driven by Democrats. The partisan differences in the effects of *justify* and *justify + racial* do not appear to be as sharp as they are for *racial*.

13 Figures 9 and 10 in the Appendix show the effect of the rebuttals on the target's feeling thermometer score and the attacker's feeling thermometer score, respectively. The figures appear to be mirror images of each other, with each rebuttal either helping the target and hurting the attacker or having no effect on either. This suggests that the effects shown in Figure 3 are not driven by evaluations of only the target or the attacker. Evaluations of both are moving in the same direction (either pro-target or no change).

14 Given 245 respondents assigned across eight conditions, the sample size per cell is about 31. Assuming a sample size of 31 per cell, an effect size of 0.15, standard deviation of 0.2, and alpha of 0.05, the power is calculated to be 83%.

15 Under these conditions, the power of the test is 86%.

16 Figure 11 in the Appendix shows the effect of the rebuttals among the racially moderate, defined as respondents in the middle third ($0.5 < RR < 0.6875$) of the racial resentment distribution ($n = 126$). The results are similar to the results for the overall sample shown in Figure 3.

References

- Ansolabehere S and Iyengar S (1995) *Going Negative: How Political Advertisements Shrink and Polarize the Electorate*. New York: Free Press.
- Bailenson JN, Iyengar S, Yee N and Collins NA (2008) Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly* 72, 935–961.
- Berinsky AJ, Huber GA and Lenz GS (2012) Evaluation online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* 20, 351–368.
- Chong D and Druckman JN (2007) Framing public opinion in competitive democracies. *American Political Science Review* 101, 637–655.
- Chong D and Druckman JN (2011) Identifying frames in political News. In Bucy EP and Holbert RL (eds), *The Sourcebook for Political Communication Research: Methods, Measures, and Analytical Techniques*. New York: Routledge, 238–267.
- Citrin J, Green DP and Sears DO (1990) White reactions to black candidates: when does race matter? *Public Opinion Quarterly* 54, 74–96.
- Cooper C (2006) Mining midterms for clues to '08 field. *Wall Street Journal*, October 24.
- Edsall TB and Edsall MD (1991) *Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics*. New York: W.W. Norton.
- Engelhardt AM (2019) Trumped by race: explanations for race's influence on whites' votes in 2016. *Quarterly Journal of Political Science* 14, 313–328.
- Frederick KA and Jeffries JL (2009) A study in African American candidates for high-profile statewide office. *Journal of Black Studies* 39, 689–718.
- Geer J (2006) *In Defense of Negativity: Attack Ads in Presidential Campaigns*. Chicago: University of Chicago Press.
- Gerber AS, Gimpel JG, Green DP and Shaw DR (2011) How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *American Political Science Review* 105, 135–150.
- Glaser JM (1996) *Race, Campaign Politics, and Realignment in the South*. New Haven, CT: Yale University Press.

- Goldman SK and Mutz DC** (2014) *The Obama Effect: How the 2008 Campaign Changed White Attitudes*. New York: Russell Sage Foundation.
- Hajnal Z** (2007) *Changing White Attitudes toward Black Political Leadership*. New York: Cambridge University Press.
- Hamilton C** (1977) Deracialization: examination of a political strategy. *First World*, March/April, 3–5.
- Hurwitz J and Peffley M** (2005) Playing the race card in the post-Willie Horton era: the impact of racialized code words on support for punitive crime policy. *Public Opinion Quarterly* **69**, 99–112.
- Hutchings VL, Walton Jr H and Benjamin A** (2010) The impact of explicit racial cues on gender differences in support for confederate symbols and partisanship. *Journal of Politics* **72**, 1175–1188.
- Jacobs J** (2018) DeSantis warns Florida not to “monkey this up,” and many hear a racist dog whistle. *New York Times*, August 28.
- Jardina A** (2019) *White Identity Politics*. New York: Cambridge University Press.
- Jeffries JL and Jones CE** (2006) Blacks who run for governor and the U.S. senate: an examination of their candidacies. *Negro Educational Review* **57**, 243–265.
- Johnston R, Hagen MG and Jamieson KH** (2004) *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge: Cambridge University Press.
- Kinder DR and Dale-Riddle A** (2012) *The End of Race? Obama, 2008, and Racial Politics in America*. New Haven, CT: Yale University Press.
- Kinder DR and Sanders L** (1996) *Divided by Color*. Chicago: University of Chicago Press.
- Krupnikov Y and Piston S** (2015) Accentuating the negative: candidate race and campaign strategy. *Political Communication* **32**, 152–173.
- Kuklinski JH and Hurley NL** (1994) On hearing and interpreting political messages: a cautionary tale of citizen cue-taking. *Journal of Politics* **56**, 729–751.
- Lau RR and Rovner IB** (2009) Negative campaigning. *Annual Review of Political Science* **12**, 285–306.
- McGraw KM** (1991) Managing blame: an experimental test of the effects of political accounts. *American Political Science Review* **85**, 1133–1157.
- McIlwain CD and Caliendo SM** (2011) *Race Appeal: How Candidates Invoke Race in U.S. Political Campaigns*. Philadelphia: Temple University Press.
- Mendelberg T** (2001) *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton, NJ: Princeton University Press.
- Mooney M** (2008) Obama aid concedes “dollar bill” remark referred to his race. *ABC News*, August 1. Available at <https://abcnews.go.com/GMA/Politics/story?id=5495348&page=1>.
- Nelson TE, Clawson RA and Oxley ZM** (1997) Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review* **91**, 567–583.
- Nelson T and Kinder DR** (1996) Issue frames and group-centrism in American public opinion. *Journal of Politics* **58**, 1055–1078.
- Nelson TE, Sanbonmatsu K and McClerking HK** (2007) Playing a different race card: elite influence on perceptions of racism. *Journal of Politics* **69**, 416–429.
- Perry HL** (1991) Deracialization as an analytical construct in American politics. *Urban Affairs Quarterly* **27**, 181–191.
- Pew** (2017) *The Partisan Divide on Political Values Grows Even Wider*. Washington, DC: Pew Research Center, October 5.
- Piston S, Krupnikov Y, Milita K and Ryan JB** (2018) Clear as black and white: the effects of ambiguous rhetoric depend on candidate race. *Journal of Politics* **80**, 662–674.
- Skoneki M** (2018) President trump calls Andrew Gillum a “stone cold thief” on fox news. *Orlando Sentinel*, October 30.
- Sniderman PM and Theriault SM** (2004) The structure of political argument and the logic of issue framing. In Saris WE and Sniderman PM (eds), *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*. Princeton, NJ: Princeton University Press, pp. 133–165.
- Sonenshein RJ** (1990) Can black candidates win statewide elections? *Political Science Quarterly* **105**, 219–241.
- Sterling A** (2009) Huckabee calls criticisms over clemency “disgusting”. *CBS News*, December 1. Available at <http://www.cbsnews.com/news/huckabee-calls-criticisms-over-clemency-disgusting/>.
- Stout CT** (2015) *Bringing Race Back In: Black Politicians, Deracialization, and Voting Behavior in the Age of Obama*. Charlottesville: University of Virginia Press.

- Tesler M** (2012) The spillover of racialization into health care: how president Obama polarized public opinion by racial attitudes and race. *American Journal of Political Science* **56**, 690–704.
- Tesler M and Sears DO** (2010) *Obama's Race: The 2008 Election and the Dream of a Post-Racial America*. Chicago: University of Chicago Press.
- Tokeshi M and Mendelberg T** (2015) Countering implicit appeals: Which strategies work? *Political Communication* **32**, 648–672.
- Valentino NA, Hutchings VL and White IK** (2002) Cues that matter: how political ads prime racial attitudes during campaigns. *American Political Science Review* **96**, 75–90.
- Valentino NA, Neuner FG and Vandenbroek LM** (2017) The changing norms of racial political rhetoric and the end of racial priming. *Journal of Politics* **80**, 757–771.
- Weaver VM** (2012) The electoral consequences of skin color: the “hidden” side of race in politics. *Political Behavior* **34**, 159–192.
- White IK** (2007) When race matters and when it doesn't: racial group differences in response to racial cues. *American Political Science Review* **101**, 339–354.