Compound Political Identity: How Partisan and Racial Identities Overlap and Reinforce

Sean J. Westwood[1] and Erik Peterson[2]

[1]Dartmouth College
[2]Texas A&M University

July 8, 2019

Abstract

The standard view conceives of partisan identity as a combination of other social identities that have sorted into alignment over a long time horizon. We develop a new and simpler theory of partisanship that responds to events in real time and does not depend on sorting. In our model of compound political identity, partisanship and race—the major American social cleavage—are so inseparable in the minds of citizens that events which independently trigger one identity also activate the other. We find support for this in three behavioral game experiments with 5,496 respondents. These studies show that changes in affect towards a racial out-group also shift behavioral and attitudinal measures of out-party affect. Conversely, changes in partisan affect spill into out-race affect. Our results suggest that negative partisan affect partially stems from spillovers from racial affect, and that the growth of partisan hostility reinforces racial hostility.

Word count: 7,711

Foundational explanations for mass political behavior emphasize social group attachments. The large portion of the public that does not engage with politics on ideological terms can instead rely on others that share their race, class, or religion (Berelson, Lazarsfeld, and McPhee, 1954; Lipset, 1960; Converse, 1964; Conover, 1988; Dawson, 1994; Achen and Bartels, 2016; Kinder and Kalmoe, 2017). Even political parties draw their relevance from the set of social group attachments they represent (Campbell et al., 1960; Greene, 1999; Green, Palmquist, and Schickler, 2002).

In an era of elevated out-party hostility, attention has turned towards how the intersection of multiple identities influences partisanship (Mason, 2018; Klar, 2018; Valentino and Zhirkov, 2018; Levendusky, 2018). Evidence shows that social sorting—the years long process that aligns an individual's social groups with their partisanship—amplifies partisan affective polarization for the subset of the electorate that is sorted (Mason, 2015; Mason and Wronski, 2018; Robison and Moskowitz, 2019) and that those who strongly associate the opposing political party with other social out-groups also hold more negative views of the other party (Valentino and Zhirkov, 2018; Ahler and Sood, 2018). While these theories are important advances, their focus on long-term patterns of alignment between different social groups means they cannot explain how individual interactions or stimuli that raise the salience of a social group identity are processed or understood by Americans.

In this paper we introduce and experimentally test a new theory of the relationship between race and partisanship. We focus on race and partisanship because of the deep historical connection between these identities (e.g., Carmines and Stimson, 1989; Valentino and Sears, 2005; Tesler, 2013; White, Laird, and Allen, 2014; Kuo, Malhotra, and Mo, 2017; Kuziemko and Washington, 2018; Valentino and Zhirkov, 2018). Our theory—*compound political identities*—argues that partisanship and race are so enmeshed in the minds of citizens that experiences which involve only one of the two groups affect evaluations and behavior towards both. Our theory draws on insights from cognitive psychology research on spreading activation in memory (Collins and Loftus, 1975; Wyer Jr and Srull, 2014; Lodge and

Taber, 2013; Morris et al., 2003). In the terms of spreading activation, partisanship and other connected groups should be thought of as nodes that are separated by some distance in memory. When this distance is large the two nodes remain separate and can be treated as distinct identities. However, when this mental distance between identities is small (based on socialization, learned experience and other factors) an explicit reference that stimulates one node will also trigger other, proximate nodes. Moreover, when an experience triggers consideration of a group to which a person does not belong, groups to which the person does belong are suppressed while other out-groups are made more salient (Lodge and Taber, 2013). The key difference between our theory and prior discussion of spreading activation is that we posit a close linkage (i.e., a short distance in memory) between race and partisanship among the entire public, whereas existing research generally asserts that relationships between groups or other political concepts vary at the individual-level.

Our theory has several advantages. First, it can explain rapid shifts in out-group affect, something missing from prior studies that consider years-long shifts in the relationship between race and party. Second, our model assumes that connections between race and partisanship are not contingent on sorting, but simply require recognition that race and partisanship are closely related. Finally, previous work largely focuses on explaining the path from race to partisanship, while the model we develop, in contrast, allows for a bidirectional relationship—from race to partisanship and from partisanship to race.

To test our model, we develop and deploy a novel experimental framework that allows us to trigger negative and positive experiences with designated out-group members. We use this to measure the direct effect of this experience on views towards the out-group emphasized in this interaction and also the spillover into affect towards other groups that were not part of the experimental treatment. Using this new approach, we find that views of partisan and racial out-groups are inextricably connected and subject to rapid updating with minimal cognition. Shifts in racial affect subsequently move partisan affect, and these shifts in affect also lead to changes in behavior towards out-partisans. Conversely, shifts in partisan affect

2

also alter an individual's views of racial out-groups.

Our new theoretical perspective and evidence extend current research in important ways. Most broadly, the rapid and interlinked shifts in out-group affect revealed in our experiments support an instinctive view of group-based political engagement that research in American politics has hinted at for decades (e.g., Converse, 1964; Green, Palmquist, and Schickler, 2002; Huddy, Mason, and Aarøe, 2015; Achen and Bartels, 2016). Beyond this, the influence of partisan affect on racial attitudes we identify substantially expands the implications of recent work on affective partisan polarization (for a review see Iyengar et al., 2019). Our results suggest that increasingly negative out-party affect is at least partially a consequence of spillovers from other identities that are more commonly triggered in daily interactions. Many are not discriminated against because of party, but are disadvantaged because of race or other social identities. It is likely, our results suggest, that these more common interactions contribute to growing partisan animus. Simultaneously, these findings suggest that the growth of partisan hostility reinforces racial hostility.

This paper proceeds as follows. First, we develop our theoretical model. Second, we introduce our experimental paradigm. Third, we present three experiments that offer strong evidence in support of our theory.

Social Identities in Politics

The dominant model of partisan identity—derived from social identity theory (Tajfel, 1970; Greene, 1999; Huddy, 2001; Mason, 2018)—suggests that partisanship exists as a construct that is distinct from other social identities. According to this theory, individuals instinctively form attachments which produce in-group favoritism and out-group antipathy. This pattern occurs based on "minimal groups," arbitrary researcher-generated divisions, and also due to real-world social cleavages such as race and class. These attachments enable a public with limited sophistication and ideological understanding to make sense of politics (Campbell

et al., 1960; Conover, 1988; Dawson, 1994; Green, Palmquist, and Schickler, 2002; Tesler, 2016; Achen and Bartels, 2016). The influence of partisan and other social identities extends into how people interact with out-group members in both political and non-political situations (Tajfel, 1970; Iyengar and Westwood, 2015).

Recent applications of social identity theory in politics extend in two directions. One line of research demonstrates that partisan identities have become increasingly salient in contemporary politics. Iyengar, Sood, and Lelkes (2012) document a substantial rise in affective polarization, the divide between the positive feelings of partisans toward the political party they identify with and their negative feelings toward the party they do not, over the past several decades (see also, Hetherington and Rudolph, 2015; Mason, 2015; Abramowitz and Webster, 2016). Whether assessed using self-reported, implicit, or behavioral indicators of partisan sentiment, negative affect toward the opposing party has reached unprecedented levels in contemporary politics (McConnell et al., 2018; Carlin and Love, 2018; Theodoridis, 2017; Iyengar and Westwood, 2015; Huddy, Mason, and Aarøe, 2015).

Another strand of research brings new attention to the relationships between social identities. "Socially sorted" individuals with multiple identities that align with their partisanship dislike their political opponents more than others with cross-cutting identities (Mason, 2018, 2016; Levendusky, 2018; Robison and Moskowitz, 2019). This pattern does not occur in all circumstances as Klar (2018) isolates conditions in which a shared identity (gender) magnifies differences between members of different political parties. Margolis (2018), Egan (2018) and Davenport (2016) reverse the canonical ordering of the social group-political party relationship, showing that political views may influence the adoption of some social identities. These studies indicate that social identities and partisanship are interconnected, but also highlight the need for further consideration of exactly how various social identities relate.

The State of the Evidence on the Relationship Between Race and Partisanship

Our theoretical discussion and empirical tests focus on the relationship between two identities: partisanship and race. This builds on an extensive literature establishing the central importance of these identities for understanding American politics (Campbell et al., 1960; Carmines and Stimson, 1989; Dawson, 1994; Kinder and Sanders, 1996; Green, Palmquist, and Schickler, 2002; Hutchings and Valentino, 2004; Valentino and Sears, 2005; Tesler, 2013; White, Laird, and Allen, 2014; Kuo, Malhotra, and Mo, 2017; Kuziemko and Washington, 2018). Existing work makes three points relevant to our present focus.

First, partisanship and race are closely linked, both in actuality and perceptually. Racial and partisan identities have become more closely associated since the 1970s (Carmines and Stimson, 1989; Green, Palmquist, and Schickler, 2002; Tesler, 2016; Kuziemko and Washington, 2018). Perceptually, Ahler and Sood (2018) show that African Americans are regarded as prototypical Democrats, so much so that the public dramatically overestimates the share of the Democratic party that is black. Whites are strongly associated with the Republican party in open-ended survey responses (Rothschild et al., 2019; Halpern and Rodriguez, 2018). Valentino and Zhirkov (2018) establish that the perceptual overlap between race and partisanship is widespread among the public and term it a "race-party schema."

Second, the causal relationship linking attitudes related to race with partisanship runs primarily in one direction, with racial identity and attitudes toward racial groups influencing the development of partisanship and views of the political parties. In discussing the formation of partisanship, Green, Palmquist, and Schickler (2002, 109) note that "based on their understanding of which groups support each party and their own affinity for these groups, many citizens come to see themselves as members of partisan groups" (see also, Campbell et al., 1960). This perspective conceives of partisan identity as the product of affect toward a political party's constituent groups, with racial groups among the most important (Converse,

5

1964). Indeed, the association between partisanship and race is posited to arise during political socialization (Valentino and Sears, 2005). The centrality of racial attitudes in the formation of partisanship (Sears and Funk, 1999) and their subsequent stability (e.g., Schuman et al., 1997; Goldman and Hopkins, ND) limits opportunities for a recursive process by which partisanship influences racial attitudes (e.g., Moskowitz and Robison 2019, but see Engelhardt, 2018).

Third, after an understanding of the relationship between race and partisanship is established, citizens are slow to update their understanding of these associations. Several examples illustrate this point. The passage of high-profile civil rights legislation prompted the public to associate the Democratic Party with African-Americans (Carmines and Stimson, 1989; Green, Palmquist, and Schickler, 2002; Kuziemko and Washington, 2018). While this led white southerners to shift to the Republican party, the crystallized nature of partisan attachments meant this process took decades (Valentino and Sears, 2005; Schickler, 2016; Mason, 2018). Likewise, the social sorting that brought partisan and racial identities into greater alignment in Mason and Wronski (2018) spans from 1972 to 2016. More recently, Barack Obama's emergence as the Democratic presidential nominee racialized partisan attitudes during his presidency (Tesler, 2013, 2016). Finally, considering social out-groups more broadly, those with polarized views of social groups related to a political party develop elevated levels of affective partisan polarization several years later (Robison and Moskowitz, 2019). These studies show a malleable association between racial and partisan attitudes, but make clear that change happens slowly, even when aided by elite rhetoric emphasizing shifts in a political party's reputation (Carmines and Stimson, 1989; Kinder and Sanders, 1996).

A Theory of Compound Political Identity

While continuing this focus on the relationship between racial and partisan identity discussed in earlier work, we draw on research from cognitive psychology to propose a model for how

these two identities relate and influence out-group affect. We start using a key observation from the "spreading activation" literature in cognitive psychology (Collins and Loftus, 1975; Quillian, 1967): related concepts in memory are activated together as a single unit. In the spreading activation model groups are defined in memory as nodes connected by varying distances. When the distances between two nodes are large, activating one node is unlikely to trigger another. However, as the distance between two nodes shrinks, activating one of them is likely to cascade activation to the other, proximate node. For instance, evoking the concept "fruit" activates nodes related to both "apple" and "pear" (Collins and Loftus, 1975).

Moving to the political realm, we argue this process has implications for how social groups are treated in memory. In particular, in the minds of Americans partisanship is systematically associated with race and that the proximity of race and partisanship in memory is sufficiently small to bind the two identities together in a compound political identity. This will lead affect towards out-groups on these dimensions to behave as a single unit (Wyer Jr and Srull, 2014). More specifically, this implies that activating racial identity (partisan identity) will also activate partisan identity (racial identity).

Our model has three important differences from the canonical models detailed in the prior section. First, unlike models of the relationship between race and partisanship that predict movement over years, our model predicts much more rapid change. This difference is because current models explain the relationship between race and party as a function of ideological alignment between these groups and the subsequent sorting of the American public (Valentino and Sears, 2005; Schickler, 2016; Mason, 2018). Our theory, however, is focused on the relationship between partisan affect and racial affect. Group attachments, and the affect that stems from them, are often presented as instinctive alternatives to more effortful political engagement that relies on cognition (Berelson, Lazarsfeld, and McPhee, 1954; Campbell et al., 1960). Achen and Bartels (2016, 228) argue that group attachments "transcend thinking" to contrast them with the cognitive nature of ideological considerations

that underpin slow and deliberate sorting (see also, Mason, 2018; Kinder and Kalmoe, 2017; Huddy, Mason, and Aarøe, 2015). Affect, unlike ideology, is implicit in nature (Kam, 2007; Mo, 2015; Iyengar and Westwood, 2015; Kinder and Ryan, 2017; Theodoridis, 2017) such that upon encountering a group-related stimulus, individuals instinctively arrive at an affective response within hundreds of milliseconds and prior to any cognitive assessments that might occur (Lodge and Taber, 2013; Ryan, 2017).

Second, our model assumes that connections between race and partisanship are not contingent on sorting but simply public perception that race and partisanship are closely related. Sorting unarguably leads to affective polarization (Mason, 2018, 2015), but our model is premised on a perceived relationship between race and partisanship and not sorted group memberships. In our model a white Democrat who sees race and partisanship as related should respond similarly to a racial treatment as a white Republican who also sees race and partisanship as related. Spillovers between groups should, our theory argues, be a function of perceived distance between groups in memory and not more complex measures of alignment such as sorting. This leads us to expect the sorted and unsorted to react similarly to stimuli.

Finally, our model predicts that spillovers are bidirectional. A wealth of earlier research identifies this pairing as having a broad association among the public in which discussions of partisanship implicitly invoke race, as well as the reverse (e.g., Valentino and Zhirkov, 2018; Ahler and Sood, 2018; Rothschild et al., 2019). Because race and partisanship are proximate in memory, triggering either should prompt changes in both.

An Experimental Design to Study Dynamic Group Affect

We use a new experimental design to assess the dynamic structure of out-group affect. While the next section fills in the specifics of this design, here we introduce the overall concept and establish its links to our theoretical motivation. We randomize the quality (i.e., positive or

negative) of an individual's interpersonal interaction with a set of out-group members. The first quantity of interest to emerge from this design is the *direct effect* of the encounter on an individual's views of the out-group, which we measure with behavioral or survey measures. We also subsequently measure their views or behavior toward other social groups (i.e., out-groups not emphasized in the interaction). This allows us to estimate a second quantity, the *spillover effect* of this interpersonal encounter on the participant's behavior or affect toward these other social groups.

Figure 1: Experimental Design to Study Dynamic Out-Group Affect
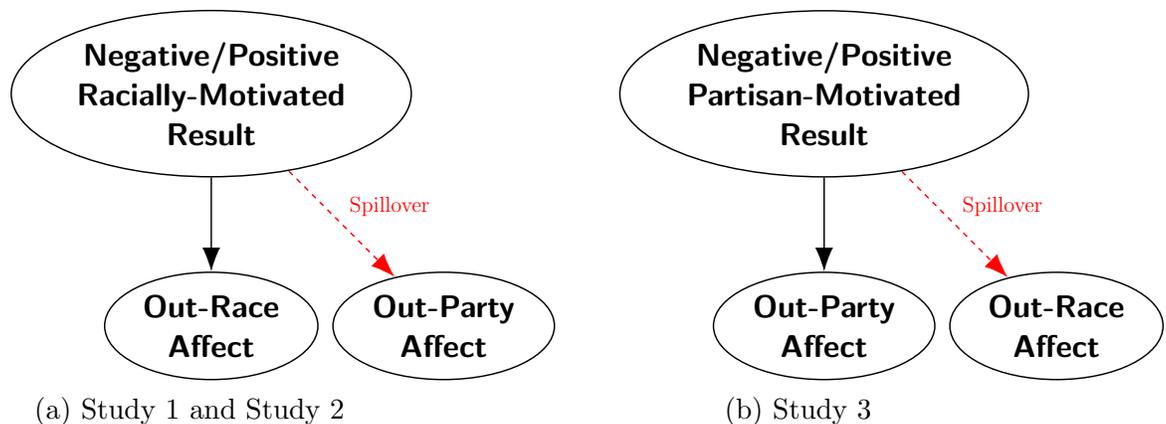

(a) Study 1 and Study 2　　　　　　　　(b) Study 3

Figure 1 shows how this design can be used to study the dynamic structure of partisan and racial affect. In each case, an initial shift in affect occurs due to an interpersonal interaction that is randomly assigned to be a positive or negative experience. In the left panel, this interaction is with a member of a different race. In the right panel, the interaction is with a member of a different political party. The solid lines indicate the direct effect of the interaction on out-group affect toward the group identity emphasized in the encounter (i.e., racial affect for the out-race treatment and partisan affect for the out-party treatment). The dashed lines indicate the spillover into affect toward the group identity not emphasized during the interaction (i.e, partisan affect for the out-race treatment, racial affect for the out-party treatment).

This design is uniquely suited to distinguish the competing theoretical motivations in-

troduced earlier in the paper. Spillover occurs when shifts in affect toward one out-group also generate changes toward a different but related out-group. Put another way, we should only observe changes in partisan affect from the racial treatment in Panel A, or the reverse pattern in Panel B, if participants simultaneously update out-partisan affect when updating out-racial affect. Evidence of this form would support our theory of compound political identity.

Delivering Treatments In a Behavioral Game

This section introduces the elements used to operationalize our experimental paradigm. We manipulate the quality of an individual's interaction with an out-group member with a heavily modified trust game (Berg, Dickhaut, and McCabe, 1995). These games are commonly used to examine incentivized, behavioral preferences for out-group cooperation by measuring the amount of money players allocate to others different from themselves on traits such as race, gender, or partisanship (Berg, Dickhaut, and McCabe, 1995; Fershtman and Gneezy, 2001; Habyarimana et al., 2007; Whitt and Wilson, 2007; Carlin and Love, 2018).

We depart from prior work by using the trust game setup to *administer* our treatment. More specifically, we manipulate the valence of a participant's interaction with a member of an opposing social group in the game. We inform participants that they are playing with an out-group member (i.e., someone from a different race or political party) and then randomize whether they have a positive experience with the opposing player (i.e., receive a generous monetary allocation from them) or a negative experience (i.e., receive no money).

Typically, Player 1 in the trust game receives a cash allocation and is instructed to give "some, all, or none" of the money to a second player. The player is also told that the researchers will *triple* any amount given to Player 2 and that this other player could, at his or her discretion, return some, all, or none of the money back to Player 1. The implication is this: the more Player 1 expects reciprocity from Player 2, the more money they should allocate to them.

Our games use this basic framework, but alter it in several ways. First, participants—when being treated—take the role of Player 2. This means they observe the allocation another player makes to them. Second, participants interact with a contrived opponent who offers allocations based on a pre-determined script. Third, we assign demographic traits to this opponent, ensuring they represent an out-group (i.e., differ from the participant with respect to either race or partisanship).

During each round, participants were told that the other player saw a demographic profile that revealed their race, age, gender, income and partisanship (Figure 2). They also saw a demographic profile of the other player displaying these same variables (Figure 2b). In Study 1, participants always played against a member of the opposing race and saw no partisan information. In Study 2, participants played against an out-race opponent, but were provided with information about this player's partisanship, which was randomly assigned at the participant-level as either a Democrat or Republican (i.e., participants played against someone of the same partisanship across all three rounds). The opponent's other profile attributes were randomly assigned in each round of play.[1]   Finally, in Study 3, participants always played against a member of the opposing political party, and the race of the opposing player was randomly assigned at the participant level (i.e., either Black or White) and remained the same in every round. With these modifications, the game unfolded like a standard trust game. Participants read instructions, saw three example rounds, and completed two comprehension questions.[2]

In an essential piece of the design, participants were randomly assigned to receive either a positive or negative result in the game. Those in the positive result condition received

---

[1]Gender was either 'Male' or 'Female', age was drawn between 25 and 35, and income was drawn from four brackets: '$30,000 - $39,999', '$40,000 - $49,999', '$50,000 - $59,999', and '$60,000 - $69,999.'

[2]If they missed a comprehension question, participants were given the answer and asked it again. Those failing the questions three times were removed from the survey.

11

allocations of \$8, \$7 and \$8 (tripled to \$24, \$21 and \$24). This greatly exceeds the typical allocation of around \$4.50 per round made in previous studies (Iyengar and Westwood, 2015; Carlin and Love, 2018). However, in a dose-response pre-test discussed in the supporting materials, we found that this average value is perceived by participants as a "negative" result. We therefore use larger values in the positive conditions to ensure participants perceive favorable intent, a pattern confirmed in manipulation checks embedded in these studies. In contrast to this positive group, those in the negative result condition were not allocated anything in each of the three rounds. Following each round, participants were told why the other player made their allocation, and for both the positive and negative result conditions, the explanation emphasized the social identity focused on in that study. This represents a strong treatment and is designed to ensure the treatment is perceived as intended. For both the positive and negative result conditions we reported the same reasoning. For Study 1 and Study 2, this was race (round 1), race and income (round 2), and race (round 3). For Study 3, it was partisanship (round 1), partisanship and income (round 2), and partisanship (round 3).[3] At the end of the game, participants saw a summary of the results for each round.

Finally, we measured out-group affect with a behavioral or attitudinal measure depending on the study. In Study 1, which used a behavioral measure, participants continued in the game as Player 1 and made allocations to four other players, allowing us to see how they interacted with out-group members when they could make the initial allocation. In Studies 2 and 3, which used attitudinal measures, participants instead reported their feelings toward several different groups in a battery of feeling thermometers. Finally, at the end of the study, participants received the value of their wallet as a cash bonus after applying a .05 multiplier.

Though stylized, this approach has several advantages. The design offers precise, randomized control over the quality of an individual's out-group interaction. It also enables participants to clearly understand the identity and motivation of the out-group member

---

[3]We added income as a reason to make our treatment less obvious.

Figure 2: Steps in the Modified Trust Game

You have been randomly
assigned to play as **Player 2** in all
three rounds.

Each Player 1 will see the following
information about you. They will use it
to decide how trustworthy you are.

Age: 26
Gender: Male
Race: White/Caucasian
Income: 50,000 – 59,999
Political Party: Democrat

(a) Step 1

Information about who you are playing
with (Player 1)

Political party:
Age: 27
Gender: Male
Income: $50,000 – 59,999
Race: Black

(b) Step 2

Results

Player 1 allocated you $8. We have
tripled this to $24.

_____

You can now return some, all or none of
this money to Player 1. Put the number
of dollars you wish to keep in the box
labeled "Player 2." Put the dollars you
wish to go to Player 1 in the box labeled
"Player 1."

Information about who you are playing
with (Player 1)

Political party:
Age: 27
Gender: Male
Income: $50,000 – 59,999
Race: Black

(c) Step 3

Play 1's reason for their allocation to
you:

- your race

(d) Step 4

Please review the game summary

Round 1:
Player 1 gave you: $8
Player 1's reason(s): your race
Earnings: $24

Round 2:
Player 1 gave you: $7
Player 1's reason(s): your race, your
Income
Earnings $21

Round 3:
Player 1 gave you: $8
Player 1's reason(s): your race
Earnings $24
Your total earnings: $69

(e) Step 5

with which they interact. Finally, although the modified trust game provides a novel delivery mechanism, this approach to shifting out-group affect via interpersonal interactions appears elsewhere in the literature on affective polarization (e.g., Kuo, Malhotra, and Mo, 2017; Suhay, Bello-Pardo, and Maurer, 2017).

Overview of Experimental Studies

We now move on to describing the results of three experiments, with a combined sample of 5,496 participants. Our first two studies test for spillover from racial affect to partisan affect. Study 1 measures this with a behavioral outcome. Study 2 uses a more common measure of affective bias, feeling thermometers and also includes several placebo measures, consideration of a group less connected to race, and a test of whether racial resentment moderates affective spillovers. Finally, study 3 reverses the direction of this test to examine potential spillover from partisanship to racial affect, again using feeling thermometer outcomes. As part of all three studies we consider whether these spillovers occur for those with sorted and the unsorted partisan identities.

Study 1: Out-Race Interactions Affect Behavior toward Out-Partisans

Based on our prior theoretical discussion, we anticipate that a positive or negative interaction with a member of a different race should also influence an individual's subsequent behavior toward members of a different political party. We test this in Study 1, which was conducted on a sample of 1,975 participants recruited via Qualtrics Panels.[4] To ensure the clarity of our treatments, we only sampled African Americans and whites and those who identified as

---

[4]The sample included quotas to approximate age and gender benchmarks from the American Community Survey.

either a Democrat or a Republican. The treatments in this study were delivered during three rounds of the modified trust game played in the manner described in the previous section. Participants learned the race of the other player, always assigned as the out-race, as well as a randomly assigned gender, income and age for each profile.

Our outcome measure is a participant's behavior toward others in additional rounds of the game played after the treatment was delivered. Using a behavioral measure in this initial experiment ensures that any spillovers we observe are consequential enough to change how subjects interact with others. This also addresses the concern that measures of out-party affect based on survey responses are expressive and reflect partisan cheer-leading (Bullock et al., 2015) or a lack of constraining social norms when assessing out-partisans in surveys (Iyengar and Westwood, 2015).

To capture behavioral responses, we observe allocations in the additional rounds of the game. Recall that in the treatment rounds, participants received allocations from a different player. In the additional rounds their role is reversed and they make the initial allocation to another player. Participants in each of the rounds were given the full allocation of $10 and a demographic profile of the contrived Player 2. In the first and second round, we offered a racial cue (white or black) and omitted partisanship. In the third and fourth round, we provided a partisan cue (Democrat or Republican) and no racial cue. As with the treatment rounds, we reported a randomly assigned gender, income, and age for each profile. Our outcome measure, the amount the participants offered to other players in these subsequent rounds, captures their incentivized, behavioral preferences for out-group cooperation.

Manipulation Check: Fairness of Interpersonal Interactions

We begin by showing that participants perceived the treatment in the intended manner. After completing the game, participants were asked, "Would you say that you were treated fairly or unfairly when playing the game?" and responded on a four-point scale that ran from "Very Fairly" (1) to "Very Unfairly" (4). Those in the negative results condition

were significantly more likely to report that they were treated unfairly (M=2.88) than their counterparts in the positive results condition (M=1.88). The mean difference in perceived unfairness between these groups is large, covering 25% of the scale's width (1 scale point; 95% Confidence Interval [.93, 1.07]).

## Effects of Out-Race Interactions on Behavior

We now move on to the substantive results of Study 1 and examine the direct effect of an interaction with an out-race member on subsequent allocations to other players of a different race. The top row of Figure 3 shows that the interpersonal interactions in the game altered behavior toward out-race players as expected. On average, those in the negative condition allocated $3.17 to the out-race player. Those in the positive condition allocated $4.97, a mean difference of -$1.80 (95% confidence interval [-2.07, -1.54]).

The second row of Figure 3 tests for a spillover from the out-race interactions into behavior toward out-partisans. When participants were treated positively by a member of the out-race in the initial rounds of the game, they subsequently behaved in a significantly more favorable manner toward members of the out-party than when they were treated poorly by an out-race member. In the negative result condition, average allocations to the out-party player were $2.72, while in the positive results condition, the allocation was $4.68. The mean difference between treatments is -$1.96 (95% confidence interval [-2.22, -1.70]). This treatment effect is of a similar magnitude to the mean difference between treatments of out-race allocations, although the group means for the out-party allocations are substantially lower than for the out-race allocations.

Movement In Behavior Toward In-Groups and Out-Groups

Figure 3 isolates behavior toward out-group members. One concern with interpreting these findings is the potential presence of a "halo" effect, in which those who had positive interactions in prior rounds of the game were simply more generous toward individuals of all

Figure 3: Behavioral Evidence of Spillover from Racial Treatments to Partisanship

Difference in allocations to the
out–group between negative and positive treatments
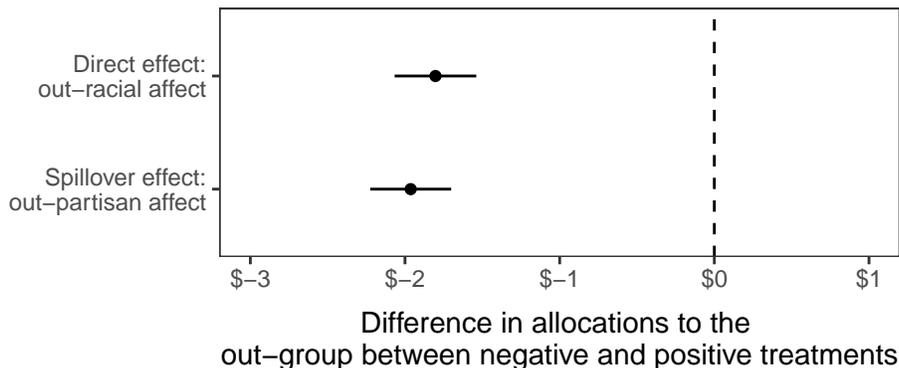
types in their later play. To address this alternative interpretation, we examine a different outcome variable, the difference in a participant's subsequent allocations to players from different parties (in-party minus out-party) and the difference in allocations between players of a difference race (in-race minus out-race). This differences out any baseline shift in allocations made to players of all types. More specifically, if participants in the positive condition are merely more positive to players of all types, we should see no significant difference on these two outcome measures.

Participants did penalize both in-race and out-race players in the negative condition, but, the penalty was larger for the out-race (mean difference = \$0.60, 95% confidence interval [\$0.37, \$0.83]). Similarly, participants offered penalty both in-party and out-party players in the negative condition, but they were again much harsher toward the out-partisan (mean difference = \$0.56, 95% confidence interval [0.32, 0.81]). The similar magnitude of these two treatment effects once again indicates a strong spillover from out-race interactions to out-party interactions.

## Spillover is Not Restricted to the Sorted

Our theory argues that spillovers occur between identities that are closely connected and simultaneously activated. Because this process occurs so quickly it does not have a cognitive component where individuals condition their responses based on the extent to which their

racial and partisan identities align. People should simply change partisan affect and racial affect jointly because the two identities are inseparable.

If this model holds, we would anticipate similar degrees of out-group spillover for those with sorted and those with cross-cutting identities. This mechanism is not dependent on any particular associations between different out-group identities; improving a white Democrat's views of the out-race (African-Americans), for instance, should lead to improved views of the out-party (Republicans). In contrast, if other perspectives explain these patterns, we should see attenuated spillover among those with a cross-cutting identity. To use the same example, if social sorting underlies this process, improving a white Democrat's views of the out-race (African-Americans) should lead to improved views of the Democratic party, not the out-party.

For this test we look at differences in spillovers for individuals who possess sorted social identities (i.e., white Republicans and black Democrats) and those possessing cross-cutting identities (i.e., white Democrats and black Republicans). Figure 4 shows that the out-race treatment affected out-party allocations on the behavioral outcome to a similar degree for those with sorted out-group identities (mean difference = $-1.95, 95% confidence interval [$-2.24, $-1.66]) and those with cross-cutting out-group identities (mean difference = $-2.02, 95% confidence interval [$-2.60, $-1.43]). An interaction testing for a difference between those with and without sorted out-group identities shows no significant difference ($\beta = \$-0.06$, 95% confidence interval [$-0.71, $0.58).

Figure 4: Sorting, Cross-cutting and Spillovers

Difference in allocations between negative
and positive treatments
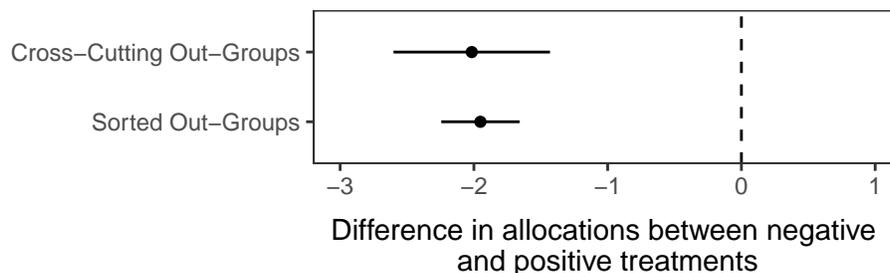
Study 1 Summary

This initial analysis offers strong evidence for our theory of compound political identity. When treated poorly by an out-race member, participants in these studies in turn behaved less charitably toward out-race players in subsequent rounds of the game. But beyond this their subsequent interactions also reveal less favorable treatment of members of the other political party, a group not explicitly invoked in their initial interaction at the start of the game. This dynamic spillover from one social out-group to another is distinct from nearly all prior work on this topic. Contrary to the stable and slow-moving linkage between racial attitudes and partisan attitudes demonstrated in prior work, we show that people instantaneously update their behavior toward members of the other political party in response to interactions with out-race members. Further, the magnitude of the spillover effect is approximately as large as the direct effect, indicating a close linkage between behavior toward out-race and out-party members. Finally, this manifests in a measure of costly behavior toward others, indicating that this pattern is not explained by various forms of expressive responding.

Study 2: Out-Race Interactions Affect Attitudes Toward Out-Partisans

Study 2 was conducted on a sample of 1,697 participants recruited via Qualtrics Panels with the same criteria as the first study. Study 2 examines the same pathway from racial to partisan affect, but it differs in several respects from Study 1. First, while the treatment is again delivered within a trust game via a positive or negative interaction with an out-race member, we also randomly assigned partisanship to these out-race players. We do so because absent this information, as in Study 1, participants might have imputed the party of the player they interacted with based on the racial cue, complicating the interpretation of our

results. This also offers a more conservative test of the theory than Study 1. Our prediction is that people automatically update their affect toward out-partisans when they update their affect toward members of a different race. If people do not update race and partisanship simultaneously and instead engage in reasoned cognition, we should see movement on out-party affect when participants play against members of a different party, but not when they face a co-partisan. Therefore, this choice to randomly assign partisanship serves to attenuate any direct effect of race on partisan affect in the half of the sample that encountered co-partisans.

Second, instead of the behavioral outcome in Study 1, we now use a feeling thermometer to measure the effects of these out-race interactions. After they had a positive or negative interaction with an out-race member, participants evaluated a number of social groups on a feeling thermometer scale. This outcome aligns our study with a number of other examinations of partisan affect that use attitudinal measures (e.g., Mason, 2018; Levendusky, 2018). Relative to game behavior, these feeling thermometers allow more direct measurement of partisan affect, the construct we aim to examine (Iyengar, Sood, and Lelkes, 2012). Because attitudinal measures are more concise than the behavioral outcomes in Study 1, their use also facilitates several robustness checks and extensions we discuss throughout the remainder of the paper.

Effects of Out-Race Interactions on Group Affect

As in Study 1, a manipulation check shows this design was interpreted as intended by participants in Study 2. Those in the positive results condition were significantly less likely to report they were treated unfairly (M=1.72) than their counterparts in the negative results condition (M=3.18). The mean difference in perceived unfairness between these groups is again large, covering 36% of the scale's width (1.46 scale points, 95% Confidence Interval [1.38, 1.54]).

Turning to the effects of the treatment on out-group affect, the top row of Figure 5 shows

that the interpersonal interactions in the game altered feeling thermometer assessments of the out-race in the expected manner. On average, those in the negative condition assessed the out-race at 52.48 points, while those in the positive condition placed the out-race at 61.86 points, a difference of -9.38 points (95% confidence interval [-11.58, -7.18]).

Figure 5: Race Treatment Spillover to Partisanship

Difference in mean feeling toward the
out–group between negative and positive treatments
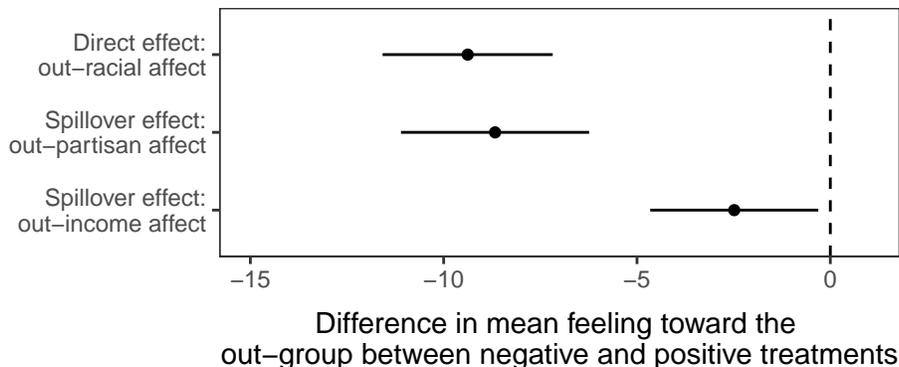
Once again, there is also a spillover from this out-race interaction into out-partisan affect. Those in the positive out-race interaction condition evaluated the out-party at 43.8 points on the feeling thermometer, while those in the negative out-race interaction condition evaluated the out-party at 35.1 points. The second row of Figure 5 displays the difference between these two groups. Relative to those with a positive out-race interaction, those in the negative condition evaluated the out-party -8.67 points (95% confidence interval [-11.10, -6.24]) lower. Once again, an out-race interaction has roughly as large an effect on partisan affect as it does on racial affect.

Internalized Spillovers or Responses to Provided Cues?

In an attempt to break the link between racial treatments and out-party affect, we randomized the partisanship of each Player 1 assigned to each participant. The purpose of this analysis is to ensure the observed changes in partisan affect stem from the racial treatment alone and do not depend on the reported partisanship of the set of Player 1s. Put another way, do partisans only update out-party attitudes when harmed/helped by an out-party

member? Or is the spillover driven by reactions to the race of the opposing player?

Our results show that participants are inattentive to the reported partisanship of Player 1 in the games. Instead, out-party attitudes are updated based on the result of the racial interaction—showing that out-groups are inseparable in respondents minds. Specifically, Table 1 shows there is never a main effect of Player 1 partisanship on out-group affect. When Player 1 was randomly assigned to be a Republican, participants were 2.62 points more positive toward the out-race (95% confidence interval [-0.43, 5.68]). Results were similar for out-party affect ($\beta = 1.94$, 95% confidence interval [-1.44, 5.31]). We also find no interaction between the (racial) treatment and the randomly assigned partisanship of Player 1 for out-race affect ($\beta = -1.51$, 95% confidence interval [-6.36, 3.33]), or out-party affect ($\beta = -1.11$, 95% confidence interval [-5.49, 3.28]). This shows that any updates to out-race and out-party affect are not an artifact of the reported partisanship of the opposing player.

We further probe the lack of an effect of the partisanship of Player 1 on out-party affect at the end of the three game trials by regressing out-group affect on co-partisanship. Table 2 shows the effect of shared partisanship of the participant and each Player 1 on out-group affect. The direct effect of the racial treatment on out-race affect remains ($\beta = 9.51$, 95% confidence interval [6.45, 12.58]), as does the spillover to out-party affect ($\beta = 11.13$, 95% confidence interval [7.75, 14.51]).

Shared partisanship between a participant and Player 1 has no effect on out-race affect ($\beta = 0.30$, 95% confidence interval [-2.76, 3.35]), out-party affect ($\beta = -1.19$, 95% confidence interval [-4.55, 2.18]), or out-income affect ($\beta = -1.68$, 95% confidence interval [-4.71, 1.34]). Moreover, there is no interaction between co-partisanship and the treatments on out-race affect ($\beta = 0.27$, 95% confidence interval [-4.11, 4.64]) or out-party affect ($\beta = -4.52$, 95% confidence interval [-9.34, 0.31]).

Table 1: Direct and Spillover Effects by Opposing Player's Partisanship

	Dependent variable:	
	Out Party Affect	Out Race Affect
	(1)	(2)
Positive Treatment	9.53***	10.19***
	(6.24, 12.82)	(7.22, 13.17)
Player 1 PID - Republican	1.94	2.62
	(−1.44, 5.31)	(−0.43, 5.68)
Participant Race - White	6.36***	10.74***
	(2.68, 10.05)	(7.41, 14.08)
Participant PID - Republican	3.55**	−0.13
	(0.98, 6.12)	(−2.45, 2.20)
Positive Treatment × Player 1 PID Republican	−1.51	−1.11
	(−6.36, 3.33)	(−5.49, 3.28)
Constant	26.84***	41.96***
	(23.00, 30.67)	(38.49, 45.43)
Observations	1,697	1,697
Note:	*p<0.05; **p<0.01; ***p<0.001	

	Out-Race Affect	Out-Party Affect
	(1)	(2)
Positive Treatment	9.51***	11.13***
	(6.45, 12.58)	(7.75, 14.51)
Player 1 and Participant Same Party	0.30	−1.19
	(−2.76, 3.35)	(−4.55, 2.18)
Participant Race - White	10.60***	7.93***
	(7.47, 13.74)	(4.47, 11.38)
Positive Treatment × Player 1 and Participant Same Party	0.27	−4.52
	(−4.11, 4.64)	(−9.34, 0.31)
Constant	43.11***	28.77***
	(39.65, 46.57)	(24.95, 32.58)
Observations	1,697	1,697
Note:		*p<0.05; **p<0.01; ***p<0.001

Movement In Behavior toward In-Groups and Out-Groups

As in Study 1, we also look for evidence of an overall halo effect. Here, we examine the difference in feeling thermometer ratings of different parties (in-party minus out-party) racial groups (in-race minus out-race). Racial polarization decreased in the positive condition relative to the negative condition (mean difference = 9.20, 95% confidence interval [6.79, 11.61]). The mean difference in affective polarization between treatments was smaller, at 2.55 (95% confidence interval [-0.62, 5.72]). While in the anticipated direction where these interactions matter more for out-groups than for in-groups, this last difference does not reach statistical significance.

Scope Conditions On Out-Group Spillover

We also test for the robustness of our results by including a second spillover group and several placebo thermometers. Our theory predicts that spillovers occur when groups are

closely connected in memory. To test this prediction we include an individual's affect toward the income out-group as a dependent measure. Income and race are less connected in the public mind (Jackman and Jackman, 1983), so based on our theory we have expectations of smaller spillover between these groups. Here, we code the in-group and out-group by family income split at the median household income in 2018. The third row of Figure 5 shows the spillover effect of the racial treatment on income, where the mean difference was -2.48 (95% confidence interval [-4.66, -0.31]). As expected, the spillover effect was stronger for tightly connected identities (race and partisanship) and weaker for the less clearly connected cleavage of income.

It is possible that people in the negative result treatment arm exited the game with a hostile attitude toward all groups, even those not referenced or connected to the attributes displayed in the game. To allay this concern, we included two neutral placebo groups of varying salience in the set of feeling thermometers: immigrants and baseball fans. There was no spillover from the treatment onto affect toward immigrants (mean difference = -1.29, 95% Confidence Interval [-3.38, 0.80]) between those in the positive result treatment arm (M = 60.00) and the negative treatment arm (M = 58.71). For baseball fans, there were also no differences (mean difference = -0.70, 95% Confidence Interval [-2.63, 1.23]) between those in the positive result treatment arm (M = 58.65) and the negative treatment arm (M = 57.95).

Spillover Is Not Moderated by Racial Resentment

It is also the spillover of race to partisanship is less general than we theorize and that it is conditional on the racism of the participant. To account for this possibility, we embedded four items from the racial resentment scale in our survey.[5] These items were measured pre-treatment and were separated from the main component of the survey with a brand recognition distraction task. We bin a participant's level of racial resentment into terciles for

---

[5]See the supporting materials for details on the questions included in this scale.

the interaction (Hainmueller, Mummolo, and Xu, 2019).[6] Using this measure, we show that racial resentment does not moderate the relationship between our treatment and either out-party or out-race affect. Instead, we find that relative to those in the first tercile (lowest racial resentment) of the racial resentment scale, those in the second ($\beta = 9.36$, 95% confidence interval [5.29, 13.44]) and third (highest racial resentment; $\beta = 4.20$, 95% confidence interval [0.03, 8.36]) terciles were more positive toward the out-party when treated well by the out-party in our experimental setup. Although, we do not test for the underlying mechanism we suspect that this result is driven by the unexpectedness of a positive interaction with a member of a derided race. This shows that the spillover from race to partisanship is not just an artifact of racism, but occurs across all participants.

Spillover Is Not Restricted to the Sorted

Finally, as in Study 1 we consider if the spillover effect is restricted to those with sorted identities or results from a more general linkage between race and partisanship. Figure 6 shows a similar pattern to study 1. Using the feeling thermometer outcome, spillover was similar for those with sorted out-group identities (mean difference = -8.09, 95% confidence interval [-11.24, -4.94]) and those with cross-cutting out-group identities (mean difference = -9.44, 95% confidence interval [-13.28, -5.61]). An interaction testing for a difference between those with sorted out-group identities and those with cross-cutting out-group identities is not significant ($\beta = -1.35$, 95% confidence interval [-6.40, 3.69]).

Study 2 Summary

Study 2 replicates our initial examination of the spillover from racial to partisan affect using a different outcome measure and treatment implementation. We once again show that racial and partisan identities proximate in memory and treated as a single unit. We also show that this relationship is not contingent on sorting, does not appear be to due to racism on the

---

[6]Results are similar when using the raw measure.

Figure 6: Sorting, Cross-cutting and Spillovers
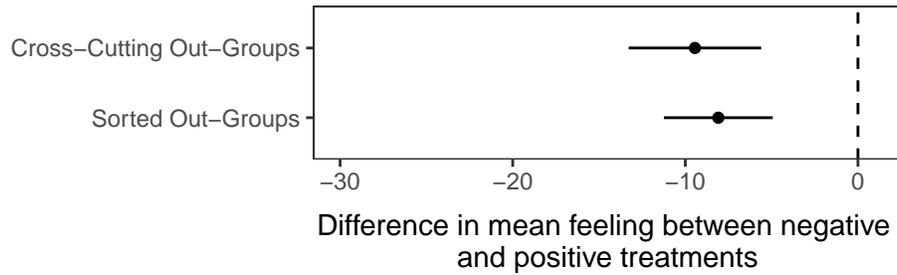
part of participants or, as tested with two placebos, just a general shift in positive affect toward all groups.

Study 3: Out-Party Interactions Affect Attitudes toward Out-Race Individuals

Studies 1 and 2 demonstrated a spillover from racial treatments into behavior and affect toward out-partisans. However, our theory of compound political identity predicts a more general relationship between salient out-groups in which the reverse pattern can also occur. Study 3 is designed to detect evidence of spillovers from partisan affect to racial affect.

Study 3 was conducted on a sample of 1,824 participants recruited via Qualtrics Panels with the same criteria as the first two studies. The treatment was again delivered over the course of a series of game rounds. Unlike in the first two studies, individuals now played against a member of a different political party. For example, Democratic participants encountered three Republican players. Like in Study 2, we filled in information about the ancillary characteristics of the other player through random draws, and in case by randomly assigning the player's race at the participant level to be either black or white.

Effects of Out-Party Interactions on Group Affect

As in the first two studies, a manipulation check demonstrates participants perceived the treatment in the manner intended. Those in the negative interpersonal interaction condition were significantly more likely to report that they were treated unfairly than those in the positive interaction condition (mean difference = 1.65, 95% Confidence Interval [1.58, 1.72]). This represents a 41% shift on the four-point scale used to record the outcome.

Figure 7 displays the direct and indirect effects of these out-party interactions on affect toward partisan and racial out-groups. We again find a direct effect of the interpersonal interactions. The average out-party affect among those assigned to have a negative interaction with an out-partisan was 32.01 points. Meanwhile, in the positive interaction condition, mean out-party affect crossed the midpoint of the scale and reached 52.21 points, a mean difference of -20.20 points (95% confidence interval [-22.59, -17.81]).

Figure 7: Partisan Treatments Spillover to Race
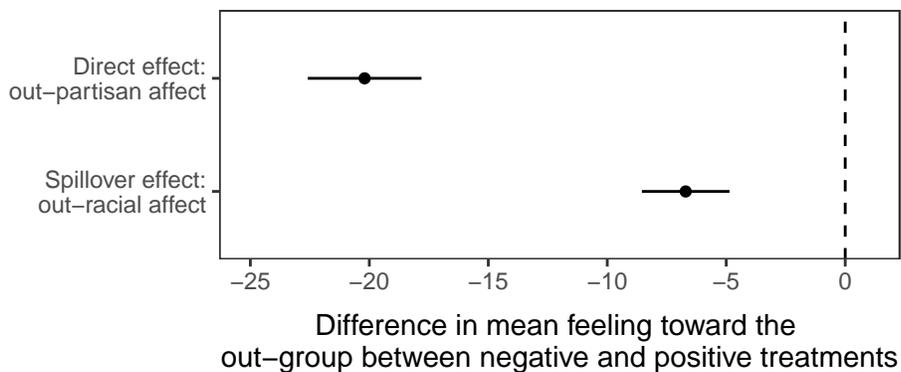
In Studies 1 and 2, we found the spillover effect of the racial treatment to partisan behavior to be approximately the same size as the direct effect of the racial treatment on racial affect. Here, we find a substantially larger direct effect (nearly twice as large) and a smaller spillover effect (77% of the spillover from race to partisanship), although we utilize a different outcome measure. Nevertheless, the spillover is still large and significant. Mean out-race affect in the negative (partisan) treatment was 55.64 points, while mean out-race affect in the positive (partisan) treatment was 62.34 points, for a difference of -6.70 points

(95% confidence interval [-8.54, -4.86]).

Movement In In-group and Out-group Affect

As with our other studies, we look for evidence of a general positivity effect. We construct racial affective polarization (in-race affect minus out-race affect) and partisan affective polarization (in-party affect minus out-party affect) measures. We find that relative to the negative treatment, the positive treatment significantly reduced racial affective polarization (mean difference = 3.64, 95% confidence interval [1.43, 5.86]) and partisan affective polarization (mean difference = 13.92, 95% confidence interval [10.86, 16.98]). Our treatment moved both in-group and out-group measures, but had a much larger effect on the out-group measures.

Spillovers Are Not Restricted to the Sorted

Finally, we again find that sorted out-groups are not necessary for a spillover (Figure 8). Indeed, although the spillover was slightly larger for sorted out-groups relative to cross-cutting out-groups, the difference is not significant in an interaction model (mean difference = -4.62, 95% confidence interval [-9.39 0.14]). The positive treatment continued to increase out-race affect relative to the negative condition for those with sorted out-groups (mean difference = -8.94, 95% confidence interval [-11.92,-5.96]) and those with cross-cutting out-groups (mean difference = -4.57, 95% confidence interval [-6.77, -2.36]).

Figure 8: Sorting, Cross-cutting and Spillovers

Difference in mean feeling between negative
and positive treatments

29
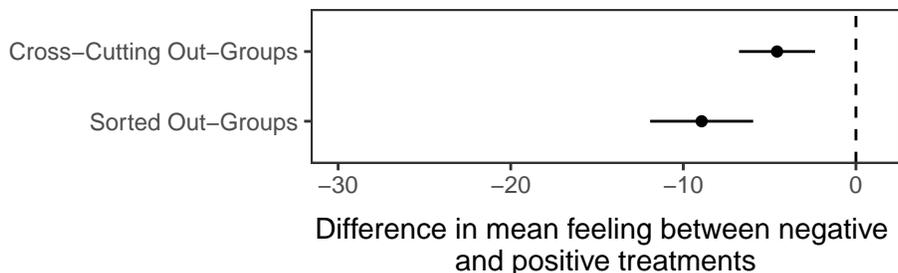
Study 3 Summary

Study 3 demonstrates that the process of spillover we identify is recursive. This is to say, beyond the race-to-party spillover that prior work emphasizes, we show that shifts in partisan affect can also spill back into out-race evaluations. The spillover from partisanship to race is considerably smaller than the spillover from race to partisanship. We suspect this more modest spillover stems in part from the social norms constraining self-reports of hostile feeling towards out-race members (Iyengar and Westwood, 2015), something that does not apply to spillovers into feelings towards members of the other political party in Study 2.

Discussion and Conclusion

In this paper we introduced a new theoretical perspective on partisanship and race that extends current research on the intersections of social identity in important ways. Our model of compound political identity explains dynamic and rapid shifts in out-group affect, an element that is missing from prior studies that consider years-long processes of shifts in the relationship between race and party. Second, our model assumes that connections between race and partisanship are not contingent on sorting, but simply require recognition of race and partisanship as closely related identities. Third, it explains responses to individual events that make social identity salient. Finally, our model predicts that spillovers are bidirectional—from race to partisanship and from partisanship to race.

We found empirical support for these theoretical claims in three experiments focused on the dynamics of out-group affect. More specifically, we show that shifts in racial affect induced by interpersonal interactions subsequently move partisan affect. The reverse pattern also occurs, as shifts in partisan affect alter racial affect. This evidence demonstrates that partisan affective polarization is not merely the consequence of a growing sense of partisanship as a social identity (Iyengar and Westwood, 2015) or a greater alignment between other social identities and partisanship (Mason, 2018). Although both contribute to this

growing affective divide, we show that partisan affective polarization is also a consequence of the non-political racial interactions of Americans. Out-race interactions rapidly spill into assessments of the other political party even for those who are not sorted. Moreover, in a finding that broadens the implications of the growing body of research on affective partisan polarization (Iyengar et al., 2019), we find that partisan hostility also contributes to racial animosity. To understand partisan affective polarization, our results suggest, requires us to consider partisanship and race not only as related groups but as inseparable in the minds of Americans.

Our results have important implications. First, they can potentially explain why partisan bias so often manifests in apolitical contexts—people have politics at mind and are unable to sever it from apolitical assessments. Second, they suggest a need to expand models of partisan affect to consider rapid spillovers from other identities that are connected to partisanship. Theorizing of partisan identity as merely a consequence of sorting among other social identities assumes a slow-moving and unresponsive relationship that is not compatible with our results. Finally, this offers a new perspective on how race and partisanship relate in American politics. At best, conceptualizing these two identities as separate is an incomplete account of how they operate. At worst, viewing these identities in a siloed manner understates the extent to which racial attitudes influence partisan attitudes and vice versa.

Our work should be assessed within a set of limitations. We focus on changes in affect toward two core social identities. Although we focus on offering initial evidence on the interconnection of race and partisanship in this paper, it is possible that the set of interconnected identities is considerably larger and that this set varies by the traits and preferences of the individual. This will be an important line of future research. Moreover, the experimental setup we deploy gives us high internal validity, but the settings are somewhat artificial. Further consideration of the real-world circumstances that can generate shifts in out-group affect of the type we study here is needed.

This paper lays a foundation for future research on the nature of group affect. First,

future work should explore the durability of these direct and spillover effects. Do these changes in affect decay back to some baseline in short order? Future work should also probe the extent to which spillover effects exist between other social identities.

References

Abramowitz, Alan I., and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of US Elections in the 21st Century." *Electoral Studies* 41: 12–22.

Achen, Christopher H., and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton University Press.

Ahler, Douglas J., and Gaurav Sood. 2018. "The Parties in Our Heads: Misperceptions About Party Composition and Their Consequences." *Journal of Politics* 80 (3): 964–981.

Berelson, Bernard, Paul Lazarsfeld, and William McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign.* University of Chicago Press.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–142.

Bullock, John G, Alan S Gerber, Seth J Hill, and Gregory A Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10: 519–578.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter.* Wiley.

Carlin, Ryan E., and Gregory J. Love. 2018. "Political Competition, Partisanship and Interpersonal Trust in Electoral Democracies." *British Journal of Political Science* 48 (1): 115–139.

Carmines, Edward G., and James A. Stimson. 1989. *Issue Evolution: Race and the Transformation of American Politics.* Princeton University Press.

Collins, Allan M., and Elizabeth F. Loftus. 1975. "A spreading-activation theory of semantic processing." *Psychological review* 82 (6): 407.

Conover, Pamela Johnston. 1988. "The Role of Social Groups in Political Thinking." *British Journal of Political Science* 18 (1): 51–76.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Its Discontents*, ed. David E. Apter. New York: The Free Press pp. 206–261.

Davenport, Lauren. 2016. "Beyond Black and White: Biracial Attitudes in Contemporary U.S. Politics." *American Political Science Review* 110 (1): 52–67.

Dawson, Michael. 1994. *Behind the Mule: Race and Class in African American Politics.* Princeton University Press.

Egan, Patrick. 2018. "Identity as a Dependent Variable: How Americans Shift Their Identities to Better Align With Their Politics." *Working Paper* .

Engelhardt, Andrew M. 2018. "Racial Attitudes through a Partisan Lens." *Working Paper* .

Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a segmented society: An experimental approach." *Quarterly Journal of Economics* pp. 351–377.

Goldman, Seth K., and Daniel J. Hopkins. ND. "Past Place, Present Prejudice: The Impact of Adolescent Racial Context on White Racial Attitudes." *Journal of Politics* Forthcoming.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds.* Yale University Press.

Greene, Steven. 1999. "Understanding Party Identification: A Social Identity Approach." *Political Psychology* 20 (2): 393–403.

Habyarimana, James, Macartan Humphreys, Daniel N Posner, and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101 (4): 709–725.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* 27 (2): 163–192.

Halpern, David, and Pedro Rodriguez. 2018. "Partisan Representations: Partisan Differences in Semantic Representations and their Role in Attitude Judgments." *Proceedings of the 40th Annual Confernece of the Cognitive Science Society* pp. 445–450.

Hetherington, Marc J., and Thomas J. Rudolph. 2015. *Why Washington Won't Work: Polarization, Political Trust and the Governing Crisis.* University of Chicago Press.

Huddy, Leonie. 2001. "From Social to Political Identity: A Critical Examination of Social Identity Theory." *Political Psychology* 22 (1): 127–156.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015. "Express Partisanship: Campaign Involvement, Political Emotion and Partisan Identity." *American Political Science Review* 109 (1): 1–17.

Hutchings, Vincent, and Nicholas Valentino. 2004. "The Centrality of Race in American Politics." *Annual Review of Political Science* 7 (1): 383–408.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76 (3): 405–431.

Iyengar, Shanto, and Sean J. Westwood. 2015. "Fear and Loathing Across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* 59 (3): 690–707.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* .

Jackman, Mary R, and Robert W Jackman. 1983. *Class awareness in the United States.* Vol. 343 Univ of California Press.

Kam, Cindy D. 2007. "Implicit Attitudes, Explicit Choices: When Subliminal Priming Predicts Candidate Preferences." *Political Behavior* 29 (3): 343–367.

Kinder, Donald R., and Lynn M. Sanders. 1996. *Divided By Color.* University of Chicago Press.

Kinder, Donald R., and Nathan P. Kalmoe. 2017. *Neither Liberal Nor Conservative: Ideological Innocence in the American Public.* University of Chicago Press.

Kinder, Donald, and Timothy Ryan. 2017. "Prejudice and Politics Re-Examined: The Political Significance of Implicit Racial Bias." *Political Science Research and Methods* 5 (2): 241–259.

Klar, Samara. 2018. "When Common Identitites Decrease Trust: An Experimental Studie of Partisan Women." *American Journal of Political Science* 62 (3): 610–622.

Kuo, Alexander, Neil Malhotra, and Cecilia Hyunjung Mo. 2017. "Social Exclusion and Political Identity: The Case of Asian American Partisanship." *Journal of Politics* 79 (1): 17–32.

Kuziemko, Ilyana, and Ebonya Washington. 2018. "Why Did the Democrats Lose the South? Bringing New Data to an Old Debate." *American Economic Review* 108 (10): 2830–2867.

Levendusky, Matthew. 2018. "Americans, Not Partisans: Can Priming American National Identity Reduce Affective Polarization?" *Journal of Politics* 80 (1): 59–70.

Lipset, Seymour. 1960. *Political Man: Social Bases of Politics.* Doubleday.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter.* Cambridge University Press.

Margolis, Michele. 2018. *From Politics to the Pews: How Partisanship and the Political Environment Shape Religious Identity.* University of Chicago Press.

Mason, Lilliana. 2015. "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59 (1): 128–145.

Mason, Lilliana. 2016. "A Cross-Cutting Calm: How Social Sorting Drives Affective Polarization." *Public Opinion Quarterly* 80 (S1): 351–377.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity.* University of Chicago Press.

Mason, Lilliana, and Julie Wronski. 2018. "One Tribe to Bind Them All: How Our Social Group Attachments Strengthen Partisanship." *Advances in Political Psychology* 39 (1): 257–277.

McConnell, Christopher, Yotam Margalit, Neil Malhotra, and Matthew Levendusky. 2018. "The Economic Consequences of Partisanship in a Polarized Era." *American Journal of Political Science* 62 (1): 5–18.

Mo, Cecilia Hyunjung. 2015. "The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters." *Political Behavior* 37 (2): 357–395.

Morris, James P, Nancy K Squires, Charles S Taber, and Milton Lodge. 2003. "Activation of political attitudes: A psychophysiological examination of the hot cognition hypothesis." *Political Psychology* 24 (4): 727–745.

Quillian, M Ross. 1967. "Word concepts: A theory and simulation of some basic semantic capabilities." *Behavioral science* 12 (5): 410–430.

Robison, Joshua, and Rachel Moskowitz. 2019. "The Group Basis of Partisan Affective Polarization." *Journal of Politics* .

Rothschild, Jacob E, Adam J Howat, Richard M Shafranek, and Ethan C Busby. 2019. "Pigeonholing partisans: Stereotypes of party supporters and partisan polarization." *Political Behavior* 41 (2): 423–443.

Ryan, Timothy. 2017. "How Do Indifferent Voters Decide? The Political Importance of Implicit Attitudes." *American Journal of Political Science* 61 (4): 892–907.

Schickler, Eric. 2016. *Racial Realignment: The Transofrmation of American Liberalism, 1932-1965.* Princeton University Press.

Schuman, Howard, Charlotte Steeh, Lawrence Bobo, and Maria Krysan. 1997. *Racial Attitudes in America: Trends and Interpretations.* Harvard University Press.

Sears, David O., and Carolyn L. Funk. 1999. "Evidence of the Long-Term Persistence of Adults Political Predispositions." *Journal of Politics* 61 (1): 1–28.

Suhay, Elizabeth, Emily Bello-Pardo, and Brianna Maurer. 2017. "The Polarizing Effects of Online Partisan Criticism: Evidence from Two Experiments." *International Journal of Press/Politics* 23 (1): 95–115.

Tajfel, Henri. 1970. "Experiments in Intergroup Discrimination." *Scientific American* 223 (5): 96–102.

Tesler, Michael. 2013. "The Return of Old-Fashioned Racism to White Americans' Partisan Preferences in the early Obama era." *Journal of Politics* 75 (1): 110–123.

Tesler, Michael. 2016. *Post-Racial or Most-Racial? Race and Politics in the Obama Era.* University of Chicago Press.

Theodoridis, Alexander G. 2017. "Me, Myself and (I), (D), or (R)? Partisanship and Political Cognition through the Lens of Implicit Identity." *Journal of Politics* 79 (4): 1253–1267.

Valentino, Nicholas A., and David O. Sears. 2005. "Old Times There Are Not Forgotten: Race and Partisan Realignment in the Contemporary South." *American Journal of Political Science* 49 (3): 672–688.

Valentino, Nicholas, and Kirill Zhirkov. 2018. "Blue is Black and Red is White? Affective Polarization and the Racialized Schemes of U.S. Party Coalitions." *Working Paper* .

White, Ismail K., Chryl N. Laird, and Troy D. Allen. 2014. "Selling Out? the Politics of Navigating Conflicts between Racial Group Interest and Self-Interest." *American Political Science Review* 108 (4): 783–800.

Whitt, Sam, and Rick K Wilson. 2007. "The dictator game, fairness and ethnicity in postwar Bosnia." *American Journal of Political Science* 51 (3): 655–668.

Wyer Jr, Robert S, and Thomas K Srull. 2014. *Advances in social cognition, Volume I: A dual process model of impression formation.* Psychology Press.