

**Drivers of Affective Polarization:  
The Relative Impact of Race, Religion, Income and Ideology**

Nicholas A. Valentino<sup>1</sup> and Kirill Zhirkov<sup>2</sup>

<sup>1</sup> University of Michigan

<sup>2</sup> University of Virginia

[Draft. Please don't cite or circulate]

## **Abstract**

Affective polarization, the growing antipathy between supporters of the Democratic Party and the Republican Party over the last 30 years, has been well documented. A compelling explanation for this trend blames the more or less simultaneous sorting of a host of group dimensions—ideology, religion, social class, and race—into the two major parties. The existing model makes three assumptions that we attempt to explore empirically: First, changes in social group coalitions for each party along each dimension changed roughly simultaneously and independently. Second, imagined party coalitions in the public mind are mostly constant. Third, perceptions of sorting on each social group dimension drive affective polarization more or less equally. Using three different measurement strategies, we find that racial sorting into the parties begins first, and has moved further than other dimensions except perhaps ideology. We also find significant variation in the imagined group coalitions of each party across the population. Finally, cognitive associations between parties, race, ideology, and religion, are quite powerfully linked to affective polarization, while objective and subjective class status have very little effect.

*Keywords:* affective polarization, ideology, race, religion, social class, sorting

Antipathy between opposing partisans in the American electorate has grown dramatically for several decades. This so called partisan affective polarization is now at its highest level since the American National Election Studies (ANES) began measuring feelings about the parties in the 1960s (Iyengar and Krupenkin 2018; Iyengar, Sood, and Lelkes 2012). On the 100-point “feeling thermometer” that the ANES uses to tap feelings about the parties, warmth expressed for members of the out-party has declined about 15 points since 1980.

Concerns about this trend include consequences like a lack of support for policy compromise, the coarsening of the national political debates (Valentino, Neuner, and Vandebroek 2018), and the degradation of basic democratic norms including the right to vote. Indeed, there is some evidence to suggest affective polarization in the mass electorate can drive ideological polarization of parties (Diermeier and Li 2019), and thwart the norms of democratic deliberation (Strickler 2018).

Partisan disdain carries important consequences outside the political sphere as well. Affective polarization has been shown to exacerbate interpersonal conflicts and discrimination in academic contexts (Rom and Musgrave 2014), trigger dehumanization toward opposing partisans (Cassese 2021; Martherus et al. 2021), and cause anxiety in families when someone marries “outside their party” (Iyengar and Westwood 2015). The extremity of these social consequences is contested (Lelkes and Westwood 2016), and can even cause public backlash against polarization in some cases. For example, people who see out-partisans being mistreated in the workplace may become less attached to their own party, effectively reducing affective polarization (Shafranek 2020). Still the bulk of the evidence to date suggests that partisan affective polarization is a potent, toxic force in American society.

So where does all this mutual dislike between Democrats and Republicans come from? Several explanations have been proposed, including ideological or issue based polarization (Bougher 2017; Gillion, Ladd, and Meredith 2020; Lelkes 2021; Orr and Huber 2019; Rogowski and Sutherland 2016), the increasing influence of negative partisan media (Lelkes, Sood, and Iyengar 2017; Levendusky 2013), and growth in the influence of moral values in partisan identity and decision-making (Garrett and Bankert 2018; also see Ryan 2017).

Several of these explanations, however, have received mixed empirical support at best, and are subject to endogeneity concerns (for a review, see Iyengar et al. 2019). For example, affective polarization has increased far faster than substantive disagreements (Fiorina, Abrams, and Pope 2006; Ansolabehere, Rodden, and Snyder 2006). In addition, warm interpersonal relations between party leaders work better to mitigate affective polarization than policy compromise (Huddy and Yair 2021). Furthermore, elite issue polarization itself might be the consequence rather than the cause of affective polarization (Diermeier and Li 2019).

In our view, the most convincing theory about the root causes of affective polarization in American politics is one that emphasizes group sorting: the growing alignment between partisanship and other social identities, such as race, religion, social class, and ideology (Mason 2018). The sorting model is theoretically rooted in the idea that individuals derive self-esteem from memberships in cherished social groups, as posited by social identity theory (Tajfel and Turner 1979). The idea is that while these social dimensions were once cross-cutting with regard to partisanship, now they are much more aligned: most religious Christians are now in one party while the non-religious are in the other, most working-class voters are now in one party, while middle- and upper-class citizens are in the other, most liberals are in one party while conservatives are in the other, and most white people are in one party while non-white citizens

are in the other. The reason partisans hate each other so much more now, the story goes, is that Americans have less in common with members of the opposite party, socially speaking, than perhaps ever before. If the hope Lipset (1960) placed in the democratically salubrious effects of “cross-cutting cleavages” was indeed well placed, their decline engenders concern. The model has amassed a significant amount of evidence in its favor: social sorting is indeed growing over time and it is correlated with attitudes and behaviors of individual partisans (Davis and Mason 2016; Mason 2015; Mason and Wronski 2018; Robison and Moskowitz 2019). This paper hopes to extend and refine the sorting model in two ways.

First, the sorting model assumes partisan coalitions along several group dimensions— party, race, religion, social class, ideology— changed simultaneously and in parallel over time, and that they therefore contribute about equally to partisan affective polarization roughly equally in the present. In other words, the model does not speculate about how each dimension of sorting might be causally related to others, and thus to affective polarization. While it is impossible to establish with certainty, it seems to us that given its structuring role in politics since the founding, that race and therefore the perceived *racial* coalitions within each party might play a leading and therefore outsized role compared to the other dimensions noted above. We will explore this possibility observationally.

Second, the sorting model assumes that the strong overlap between social cleavages and partisan coalitions is common knowledge, and therefore that there is little variation across the electorate in the imagined group coalitions of each party. For example, once a substantial proportion of non-whites come to identify with the Democratic party, the logic goes, the party is imagined to be “non-white” by most members of the public regardless of their own partisanship. We suspect, instead, that actual group sorting is a moving target, that subjective perceptions vary

from person to person, and that this variation matters a great deal for affective polarization.

Other work has begun to develop both implicit and explicit measures of individual level variation in these schemas (Ahler and Sood 2018; Claassen et al. 2021; Zhirkov and Valentino 2021). We will use them to strengthen the tests of the sorting model.

The paper, therefore, hopes to make several important contributions. We first connect real demographic partisan sorting with social psychological research on social perceptions and their political consequences. Based on this synthesis, we propose a modification of Mason’s (2018) model of affective polarization that emphasizes the importance of individual-level beliefs about social compositions of partisan coalitions that we call “group–party schemas.” To test this theory, we deploy individual level measures of group—party schemas. The general pattern of results points to racial sorting as temporally prior to and faster than sorting on other dimensions. Further, race–party schemas measured implicitly and explicitly are a powerful predictor of affective polarization in two separate studies. Schemas linking parties to religion and ideology are also important, but class returns consistently weak results. We end the paper by proposing a way to think about the causal ordering of social sorting on these various dimensions and antipathy between partisans in the United States.

### **Partisan Sorting and the “Images in Our Heads”**

Our approach begins with the view that partisanship is an aggregator of social identities translated into the realm of politics (Green, Palmquist, and Schickler 2002). In the past, American political parties represented diverse political coalitions that attracted supporters on the basis of multiple social identities such as geographic region, urban vs. rural residence, Christian religiosity, race/ethnicity, social class, and so on. This situation created multiple social divisions that were largely independent from each other. These divisions became an important source of

political stability in the United States after the Civil War. Cross-pressured voters maintained social loyalties that pulled them simultaneously towards different partisan camps (Lazarsfeld, Gaudet, and Berelson 1948). When an individual's group memberships come into alignment with partisan coalitions, however, voting behavior is much more predictable (Zuckerman, Valentino, and Zuckerman 1994).

Other classic accounts of partisanship in the United States are consistent with this story. For instance, the authors of *The American Voter* have found that whenever partisanship changes, it happens as a response to shifts in respondents' social group memberships (Campbell et al. 1960). The stability of partisanship is thus rooted in social stratification and mostly invariant psychological attachments to groups socialized at an early age (Zuckerman 2012). Partisanship acts something like an aggregator of "primary" social group memberships, such as race/ethnicity, religion, social class, and so on (Zuckerman, Valentino, and Zuckerman 1994; Green, Palmquist, and Schickler 2002). According to this view, the development of partisanship follows a two-step process. First, the individual comes to identify with primary social groups such as religion, race, class, and so on. Second, these attachments lead her to associate with the party seen as closest to those groups, and which actively distances itself from outgroups she dislikes. This view is consistent with evidence from Ostfeld (2019), who finds white Democrats feel significantly less warmly toward their party when it makes appeals to other ethnic groups, such as Latinos.

This social perspective on partisanship, with its special attention to overlapping versus crosscutting group identities, seems a necessary condition for the affective polarization we now see. However, the current approach has left one moving piece unmeasured, and that is variation across voters in the content of their party images. Therefore, it is necessary to isolate and measure variation in these cognitive connections between social groups and parties. Powerful

negative affect between partisans depends upon a clear mental image of the out-party as a home for social out-groups.

We refer to these mental images as “group–party schemas.” The term “schema” in this context denotes a “cognitive structure of organized prior knowledge” about social groups (Fiske and Linville 1980, 543; see also Conover 1988; Conover and Feldman 1984). Overall, our model implies that partisan dislike springs from a schema of the opposition party as composed of disliked out-groups, while the in-party is made up of cherished groups with which the individual identifies.

But for group–party schemas to drive affective polarization, they must vary in strength. The sorting theory implicitly specifies groups as either in or out of each party’s coalition. For example, Evangelical Christians are commonly perceived as supporters of the Republican Party. However, this alignment is rather recent, and is the result of continuous sorting over time: Southern Evangelicals were far more Democratic in their partisanship several decades ago. If the provenance of a party’s coalition is dynamic over time, then some of the effect of variation in identification with groups in that party’s coalition may actually be due to changing perceptions, rather than reality. For example, the perception of Democrats as “secular” and the Republicans as “Christian” may according to a respondent’s personal experiences, social networks, and exposure to elite cues. We therefore must develop measures to capture variation in the cognitive linkage between parties and social groups across the population.

### **The Role of Race**

Our second contribution focuses on comparing the power of group–party schemas in driving affective polarization. As mentioned above, the sorting theory assumes that different group–party alignments (religion–party, race–party, class–party, and ideology–party) have roughly equivalent

effects on affective dislike of the out-party. This assumption may be correct in expectation across political systems, on average. However, we suspect that in any given country, some group dimensions will be much more powerful than others. The reason is that some social group dimensions are more salient than others in a party's coalition, and the intensity of attitudes about groups also varies across societies. In the United States, we suspect that racialized schemas of each party carry the largest weight in driving the polarization that has taken place over the last few decades, while other group dimensions—class, religion and even perhaps ideology—might have smaller effects. We base this prediction on three core observations.

First, racial conflict has a long and well documented pedigree as a driver of the party system in America. In the late 1800s and early 1900s, racial cleavages likely prevented formation of strong labor unions and emergence of a European-style socialist movement (Lipset and Marks 2000). In the post-Reconstruction South, stoking up racial animosity among rural whites became the backbone of the single-party rule (Key 1949; Mickey 2015). Beginning in the 1950's, the struggle to end discrimination against and disenfranchisement of African Americans caused a sweeping partisan realignment (Carmines and Stimson 1989), whereby the Republican Party captured racially conservative whites in the south and, eventually, in the north as well (Valentino and Sears 2005). Black voters, in the meantime, shifted dramatically to the Democratic party. Racial divisions were then weaponized by political actors opposing income redistribution (Gilens 1999), and criminal justice reform (Valentino 1999). Racial conflict is therefore one of the cornerstones of contemporary American politics, influencing political attitudes, behaviors, and government institutions (Cramer 2020; Hutchings and Valentino 2004; McClain 2021). Indeed, many would argue racial conflict has American roots far deeper than that.

The second observation concerns the timing of affective polarization. The downward trend in out-party thermometer ratings began in the 1980s and persists to the present day (Iyengar, Sood, and Lelkes 2012). The trend, in other words, seems to have begun roughly one generation after the civil rights movement began. If race is at the heart of affective polarization, why did it take a generation to have its effect? One answer is that the partisan realignment itself was driven by generational replacement, rather than individual party switching. Only as cohorts socialized before 1960s and 70s passed away did images of the Democratic Party as non-white and the Republican Party as white become entrenched in the public mind (Osborne, Sears, and Valentino 2011).

Third, the political influence of race and racial attitudes in the United States is growing. While certainly alive and well at the turn of the 21st century (Kinder and Sanders 1996), racial identity and racial attitudes now strongly predict opinions and behaviors of Americans of all races. Since Barack Obama's victory in 2008, the debate about the centrality of race and racial conflict in American politics has intensified (Tesler 2016; Tesler and Sears 2010). The country's first African American president polarized the electorate and increased opposition to left-wing policies among racially conservative whites (Tesler 2012; Filindra and Kaplan 2016), may have increased racial resentment (Valentino and Brader 2011), and mollified the public's previous aversion to explicitly hostile racial rhetoric (Valentino, Neuner, and Vandebroek 2018).

We would argue that none of the other social cleavages highlighted by the sorting model—religiosity, class, and ideology—have the same historical provenance in structuring party coalitions. These observations point to the same prediction: when it comes to increasing affective polarization over time, racial sorting and the subsequent strengthening of racial schemas are probably at the heart of the matter. Two unique and testable empirical expectations

spring from this conjecture. First, actual partisan sorting on the basis of race should have begun earlier, and should be more powerfully associated with affective polarization over time compared to sorting on the basis of other group identities. Second, racialized schemas about partisan coalitions should be a stronger predictor of affective polarization than schemas linking parties to non-racial group identities. We test these predictions across three studies.

### **Study 1: Group Sorting and Affective Polarization over Time**

To test our first hypothesis, we employed time-series data from the American National Election Studies (ANES).<sup>1</sup> These data compared partisan sorting on the basis of four major social cleavages—race, religion, social class, and ideology—and estimated their relative importance for affective polarization in the American electorate. To record voters' characteristics on these identity dimensions, we used the following ANES items: First, each respondent is racially categorized by the interview as *White non-Hispanic*, *Black non-Hispanic*, and *Other*. Second, self-reported religious affiliation coded 1 = *Protestant or Catholic*, 0 = *All other religious groups*. Third, household income coded from 1 = *0 to 16th percentile*, to 5 = *96 to 99th percentile* as a proxy for social class. Fourth, self-assessed ideology on the standard seven-point scale from 1 = *Extremely liberal* to 7 = *Extremely conservative*. In choosing these measures, our goal was to maximize the length of the time series and comparability across election years. Self-reported ideology, for example, is available only since 1972.

To compute partisan sorting measures on the basis of these four items, we took the following steps. First, we normalized all non-binary measures—i.e., other than race and

---

<sup>1</sup> The American National Election Studies ([www.electionstudies.org](http://www.electionstudies.org)). These materials are based on work supported by the National Science Foundation under grant numbers SES 1444721, 2014-2017, the University of Michigan, and Stanford University.

religion—to run from zero to one (see Online Appendix).<sup>2</sup> The resulting individual-level scores had the same direction, with higher values corresponding to stereotypically Republican groups: white, Christian, affluent, and conservative. Second, we calculated the proportion or mean on each of these dimensions for both Democratic and Republican identifiers, including leaners. Third, we subtracted the Democratic score from the Republican score on each social group dimension to obtain a group sorting score for a given year. A sorting score of zero, therefore, indicates no difference in terms of the composition of the two parties on a given group dimension. A sorting score of one, in turn, would mean the highest possible degree of sorting on that group dimension (e.g., all Republicans are white, Christian, affluent, or conservative whereas all Democrats are non-white, non-Christian, poor, or liberal).

To measure partisan affect, we relied on differences in thermometer scores towards the Democratic Party and the Republican Party: items in which respondents reported their feelings using 100-point thermometer ratings. We calculated individual-level partisan affective polarization scores as absolute difference between the two parties' thermometer ratings. We then computed the mean affective polarization score for each election year. Beginning in 1964, the ANES asked respondents to rate how warmly they felt toward “Democrats” and “Republicans.” In 1978, the wording changed slightly, so respondents rated “the Democratic Party” and “the Republican Party.” As a result, these analyses include 23 elections over 52 years, from 1964 to 2016. Also, to ensure comparability across years, we used only the face-to-face samples in 2012 and 2016, when online samples were also available.

## Results

---

<sup>2</sup> We preserved the scales' relative intervals. For instance, the ANES income scale was from 1 (least affluent) to 5 (most affluent) with one-unit intervals. After recoding, the measure runs from 0 (least affluent) to 1 (most affluent) with 0.25 unit intervals.

We take the growth of affective polarization in America over the last several decades to be a settled empirical fact based on the work reviewed above. Our goal is to understand whether the trend is more closely linked to racial versus another dimension in the constellation of group identities proposed by the sorting model (Mason 2018). To find out, we trace racial and non-racial partisan sorting and estimate its effect on affective polarization from 1978 to 2016.

We start by examining social sorting via changes in composition of the Democratic Party and the Republican Party, including leaners as partisans and excluding pure independents, since 1952. Figure 1 presents the composition of each major party coalition with overlaid polynomial trends for a maximum of 30 election cycles in the case of race, religion, and income, and 20 cycles in the case of ideology. A casual inspection of these patterns suggests the sorting into the parties along both race and ideology are strongest. At the beginning of the time series in the first panel, both parties were nearly entirely white. As the country become more diverse, so did both parties, but the trend was far steeper for Democrats as both African Americans, new immigrants, and smaller non-white groups predominantly chose that partisan identification as they entered the electorate. In 1952, 90% of Democratic Party identifiers also were identified as white, compared to only 57% in 2016, a decline of 33 percentage points. The Republican Party changed only by 12 percentage points over the same period, from 95 to 83% white. Note that the population as a whole, according to the ANES time series, became about 21 percentage points less white over this period. The sorting of liberals and conservatives into each party has been similarly rapid, though it is difficult to compare with race precisely since the time series begins later. In 1972, the earliest year the ANES measures symbolic ideology, the mean score for Democratic identifiers was .47 on the 0–1 (liberal-conservative) self-placement scale. Note that .5 on this 0–1 scale corresponds to a 4, or “middle of the” road on the 7-point scale running from 1 = *Extremely*

*liberal* to 7 = *Extremely conservative*. In 2016, the mean ideological score for Democratic identifiers dropped to .37, corresponding to “slightly liberal.” Republicans have shifted in the other direction, obviously, from .58 (middle of the road) to .68 (slightly conservative). Substantively speaking, therefore, this is not dramatic sorting in terms of ideology.

Partisan sorting along religious lines has been less pronounced, since the entire population has become less Christian and more secular over time regardless of party (Voas and Chaves 2016; also see Baldassari and Park 2020). In 1952, 92% of Democrats and 97% of Republicans identified as Christians (Protestant or Catholic denominations versus all non-Christian groups plus secular identifiers). By 2016, only 67% of Democrats and 78% of Republicans identified as Christians. Therefore, the partisan gap in Christian identification increased only from 5% to 11% over the 64-year time series examined here. Finally, partisan sorting by income has been rather stable over time, and may even be closing recently as Democrats attract more affluent voters. The median income for both sets of partisans has remained fairly close to the middle of the income distribution for the entire country.

[Figure 1 about here]

Based on the results in Figure 1, one might expect to find that sorting along racial and ideological dimensions would have the largest impact on partisan affective polarization over time. Of course, this is not necessarily true, since rather small differences in partisan coalitions along any group dimension could contribute substantially to affective disliking of out-partisans. We begin to explore this by estimating linear bivariate relationships between sorting on the four group dimensions and partisan affective polarization. Figure 2 presents these results, now relying on the 23 elections between 1964 and 2016 where we have measures of partisan affect. The simple association between racial sorting and affective polarization is very large ( $\beta = 66.0$ ,  $r =$

.70), and statistically significant ( $p < .001$ ), even when based on only the 23 elections since the ANES began measuring party affect. The association between religious sorting and affective polarization over time seems even larger ( $\beta = 84.1, p < .05$ ), but notice there is less precision in this association ( $r = .47$ ), especially prior to 2008. As expected, the linear association between income and affective polarization is much weaker and statistically insignificant ( $\beta = 32.7, p = .47$ ), and the simple bivariate correlation is very small ( $r = .16$ ). Finally, the linear association between ideological sorting and affective polarization is slightly larger than for race ( $\beta = 78.1; p < .001$ ), as is the simple correlation ( $r = .77$ ). These results do not suggest race is obviously a larger force in racial polarization over time compared to ideology at least, but they do hint at the possibility that religion and income are not as central. We will have more to say in the conclusions about whether racial or ideological schemas likely came first, causally speaking.

[Figure 2 about here]

Taken together, these results suggest that racial and ideological partisan sorting were both pronounced and tightly linked to affective polarization over time, at least compared to religiosity and income. However, this is highly aggregated and indirect evidence for testing the central claim in this paper, that variation in group–party schemas is at the heart of affective polarization, and that changes in race–party schemas in particular are crucial. To explore that possibility, variation in group–party schemas must be measured more directly, at the individual level. We begin to make that effort in Study 2.

### **Study 2: Explicit Group–Party Schemas and Affective Polarization**

Our goal in Study 2 was to directly measure respondents’ beliefs about the compositions of the U.S. partisan coalitions on the same four cleavage dimensions explored in Study 1—race, religion, social class, and ideology. We could then estimate their relative associations with

affective polarization. In February 2016, we conducted a survey using a sample of American adults on the Amazon Mechanical Turk (MTurk) platform. After removing duplicated IP addresses and responses from outside of the United States, the final sample contained 466 observations out of 520 completed interviews. Sample demographics deviated from national parameters in expected based on previous research regarding the MTurk platform (Berinsky, Huber, and Lenz 2012). Our respondents were disproportionately college educated (53%) and male (56.9%). The sample also leaned Democratic (48.1%), and young (mean age was 36.4 years). Only 18.5% of respondents were Republicans, and 30.5% identified as independent. Finally, 76.2% of the sample identified as non-Hispanic white. While this sample was not representative of the American population, it nonetheless afforded us the ability to measure, and then compare, differences in the strength of mental associations between group and parties.

In the survey, each respondent was asked about his or her own identification with regard to party, race/ethnicity, religion, social class, and ideology. Respondents were then asked to identify the “the typical supporter” of both the Democratic Party and the Republican Party on each of the same identity dimensions. For example, respondents were asked “In terms of race, which of the following comes closest to describing the typical supporter of the [Democratic Party/Republican Party]?” Possible answers were *White/Caucasian American*, *Black/African American*, *Asian American*, *Native American*, *Hispanic/Latino American*, and *Other*. See Online Appendix for all questions and response options.

## **Results**

Table 2 presents how often respondents choose specific group identity categories as “typical” of Democrats and Republicans. Overall, there is more diversity in how respondents view the Democratic Party compared to the Republican Party. On race, almost 30% of respondents in this

relatively Democratic sample believed that the typical Democrat is black, whereas the typical Republican is seen as white by nearly all respondents. In terms of religion, more than one third of the sample guess that the typical Democrat is secular, but many also name non-Evangelical Christian denominations (such as Mainline Protestant, Catholic, and non-denominational Christian; we collapse all these into the “Other” category for the purpose of this analysis). On the other hand, more than half of the sample view the Republican Party as Evangelical, with only about 1% of respondents positing the typical member of that party is secular. With regard to social class, roughly equal shares of our respondents placed Democrats in working or lower class and middle class or rich/wealthy. On the other hand, almost 80% of our respondents placed the typical Republican in the middle class or rich/wealthy category.<sup>3</sup> Finally, there was nearly complete consensus in the sample that the typical Democrat is liberal and the typical Republican is conservative, though perceptions of the Democratic party were perhaps slightly more diverse ideologically. In summary, group–party schemas as tapped with this explicit measure reveal widespread consensus that Democrats are liberal, racially diverse, and working or lower class, while Republicans are conservative, mostly white, Evangelical, and middle class or wealthy. There was more variation about the typical Democrat on all these social group dimensions. This is an important distinction. On its face, race and ideology are the most thoroughly sorted group dimensions in the minds of respondents in this sample. The question, however, is how tightly are these perceptions linked to affect about the parties?

[Table 2 about here]

---

<sup>3</sup> Many respondents, when asked about the class status of the typical Republican, chose the “Other” option and answered “rich” or “wealthy.” In the analysis, we collapsed them with our pre-defined categories so that “working class” effectively meant “working or lower” whereas “middle class” meant “middle class or rich/wealthy.”

To answer that question, we construct a measure of the perceived match between parties and groups for each individual. We combine responses to the questions about typical partisans with a measure of the respondent's group identity. We then can calculate whether the individual is matched to one party, while also believing outgroup members are matched to the other party on the four group dimensions we have been discussing: race, religion, social class, and ideology. A matching score with three distinct values can then be created for each group dimension:  $-1$  (the respondent is matched with the Democratic Party on that group dimension),  $0$  (matched with neither party), and  $1$  (matched with the Republican Party).

Table 3 presents our coding procedure, using religion as an example. Overall, to produce a non-zero score, a respondent would perceive one of the two parties as matching his or her identity and, simultaneously, the other party as closer to the outgroup. For instance, if a Christian respondent perceives Republicans as Christian and Democrats as either secular or members of any other religious group, we assign them a matching score of  $1$  because the respondent's religious identity is matched to the group that they perceived typical of the Republican Party. Similarly, if a Christian respondent posits the typical Democrat to be Christian but the typical Republican to be secular or a member of some other religious group, they would be scored a  $-1$  because the respondent's religious identity is matched to the group that they perceive typical of the Democratic Party. The same procedure is performed for secular respondents, with the identity categories flipped. Note that respondents are considered unmatched (a score of  $0$ ) if they imagine the typical member of both parties to belong to the same social group (i.e., both Christian or both secular), regardless of their own religious identity. The resulting measure effectively captures whether a person's schema for a given party matches his or her own social

group identity. See Online Appendix for scoring on the other dimensions, which were created in the identical way.

[Table 3 about here]

Using this measure, we can compare how tightly each group–party schema is linked to affective polarization. Figure 3 presents the results graphically. The dependent variable in the reported regression models is partisan affective polarization: the Republican minus the Democratic feeling thermometer centered around the observed mean. We run OLS models for each of the four group–party measures coded –1, 0, and 1. Controls include age, gender, education as well as positions on the following issues: size of government, defense spending, environmentalism, and abortion. To interpret these figures, remember that the *Y* axis will take on negative values when the Democratic party is preferred to the Republican party. In the first panel, we see that among black respondents who believe the typical Democrat is a member of their own race, and the typical Republican is white, the Democratic Party is rated an average of 10.7 degrees more warmly than the Republicans. Among whites those who believe the typical Republican is a member of their own race and the typical Democrat is not white, the Republican party is rated 6.4 degrees more warmly than the Democrats on average. The affective difference between respondents racially matched to the Democrats versus the Republicans is therefore more than 17 degrees. Among those who do not think the typical members of the Democratic Party and the Republican Party differ by race, affective polarization is statistically indistinguishable from zero. The slope of the line connecting these three points is indicated by the regression  $\beta$  coefficients at the bottom of the figure, and is statistically significant in the case of race.

The effect of matching with a party on religion is weaker than for race, and matching with class is essentially unassociated with affective polarization in this sample. This pattern

holds after controlling for individual-level demographics and issue positions (e.g., the size of government, defense spending, environmentalism, and support for women’s reproductive rights). Of the four group–party schemas, race and ideology are again most powerfully linked to affective polarization. The difference between people who believe their own race coincides with the typical Republican and those who believe their race to be more typical of the Democratic Party is almost 18 points on the affective polarization scale. The corresponding effect of ideological matching, however, is much larger: almost 60 points. Note that this pattern would be expected if ideological group identities are primary drivers of affective polarization—but also if they are in fact the downstream consequences of racial attitudes. Why might this be the case?

[Figure 3 about here]

We have two thoughts about why the ideology–party schema is a stronger predictor of affective polarization than any other dimension including race. First, it could be the case that perceived differences in the racial coalitions of each party have, over time, come to define ideological differences. There is plenty of empirical evidence to suggest that Americans, especially whites, stereotype non-whites as liberal and whites as conservative (e.g., Berinsky et al. 2020). If racial stereotypes drive assumptions about ideology, then the effect of ideological beliefs about the party coalitions on affective polarization is itself a result of attitudes about racial groups. It is very difficult to rule out this possibility. But why would the effect of ideology be so much stronger than race using this matching technique?

This might have something to do with social desirability biases in the measure. Since the matching technique here depends on respondents’ assessments of their own identities as well as those of “typical” partisans, racialized party schemas might be underreported relative to connections between ideology and party. Respondents might be more reluctant to admit that they

view the parties as highly distinct along racial lines than along ideological lines. It would be less socially risky to report the belief that the typical Republican is conservative and the typical Democrat is liberal. We attempt to address this limitation in Study 3 by deploying an implicit measure of group–party schemas.

### **Study 3: Implicit Group–Party Schemas and Affective Polarization**

To measure group–party schemas implicitly, we modified the implicit association test or IAT (Greenwald, McGhee, and Schwartz 1998). The original IAT asks respondents to quickly classify stimuli that appear on the computer screen using pre-defined keys. One key, for example, is assigned to “positive” words, and the other to “negative” ones. The first key is also assigned to one class of objects (e.g., flowers) while the second is assigned to another (e.g., insects). These pairings assigned to a particular key are switched during the task. When a picture (flower/insect) or word (pleasant/unpleasant) are flashed on the screen, the respondent needs to press the key to which that stimulus is assigned as quickly as possible. The assumption underlying the task is that respondents will take less time to press the correct key when stereotypically related objects and affective states (e.g., flower/pleasant and insect/unpleasant) are assigned to the same key than when they are stereotypically incompatible (e.g., flower/unpleasant and insect/pleasant). The method has been used most commonly to assess automatic, implicit affective evaluations of social categories such as racial or gender groups (Nosek et al. 2007).

We modified the task to examine variability in the strength of cognitive pairing between attitude objects, like parties and social groups, rather than the implicit affect a person would hold for a single object. Instead of positive and negative words, we pair party symbols with those from various social groups to measure the strength of implicit associations between the two. So,

for example, respondents are told that one key represents the group “black” and also the group “Democrat”, while another represents the group “white” and the group “Republican.” When a picture or word representing one of these concepts is flashed on the screen, the respondent quickly chooses the key to which it corresponds. Like the standard IAT, we assume that when a key is paired with stereotypically consistent groups (white–Republican, black–Democrat) response times will be faster than when keys are paired counter-stereotypically (black–Republican, white–Democrat).

The differences in the times that respondents take to classify compatible vs. incompatible pairings of categories/attributes are known as D-scores, which measure strength of the implicit association in memory. If these cognitive pairings are weak, of course, the response times across these two pairs will be roughly equal, and the ultimate scores on the task close to zero. The theoretical maximum for D-scores is  $-2$  to  $2$  (Greenwald, Nosek, and Banaji 2003), though typically most scores fall between  $-1$  and  $1$ . In the case of the current tests, larger positive scores would indicate stronger associations between stereotypically consistent attitude objects (white–Republican, black–Democrat), while negative scores would indicate counter-stereotypical implicit associations (black–Republican, white–Democrat).

We used the same architecture to record implicit associations between the social categories of interest (race, religiosity, class, and ideology) on the one hand and the Democratic Party and the Republican Party on the other hand. For parties, we used a collection of publicly available official and unofficial symbols of the two parties: DNC and RNC logos, donkey vs. elephant, and so on. For social groups, we used word stimuli pre-tested in a separate MTurk sample. An important distinction between this study and the previous two involves the measure of religion-party schemas. Here to examine the pairing between the concepts religious/non-

religious groups and parties we used a broader definition. Instead of using the words “Christian” and “non-Christian”, we used the words “evangelical, religious, churchgoer, and worshiper” to cue religious people versus “secular, non-religious, atheist, and non-believer” to cue non-religious people. This measure allows us to see whether individuals think of the broader distinction between religious and non-religious people when imagining the parties. Still, in the U.S. context, the vast majority of highly religious people are Christians.<sup>4</sup> The IAT was implemented in Qualtrics using a JavaScript-based tool (Carpenter et al. 2019).

We collected these data in December 2019, again using Amazon MTurk to recruit respondents. Since the MTurk participant pool is known to overrepresent liberals, we oversampled moderates and conservatives. The result was a relatively balanced sample in terms of partisanship: 48.5% were Democrats and 41.2% were Republicans (including leaners). The modal age was between 25 and 34 years, 54.7% of respondents were female, 58.8% were college educated, and 75.2% identified as white.

The time and cognitive effort required to complete an IAT is substantial, so each respondent was assigned to one and only one of four group–party schema measurement tasks. Altogether, we collected 2,000 completed interviews with approximately 500 respondents in each IAT. We also measured affect about the Democratic Party and the Republican Party as well as about relevant social categories—whites, religious people, rich people, and conservatives—using standard feeling thermometers. See Online Appendix for question wording.

## Results

---

<sup>4</sup> To cue whites we used White, European American, Caucasian, Anglo. To cue non-whites we used Black, African American, Minority, Latino. To cue wealthy people we used Rich, Wealthy, Well-off, Upper class. To cue poor people we used Poor, Needy, Broke, Lower class. To cue conservatives we used Conservative, Right-wing, Rightist, Capitalist. To cue liberals we used Liberal, Left-wing, Leftist, Socialist. See Online Appendix for the full list of the IAT stimuli.

We start with the descriptive results: the empirical distributions of implicit group–party schemas measured via IAT D-scores. Figure 4 displays the empirical densities of D-scores for Democrats and Republicans (including leaners), with differences between Democrats’ and Republicans’ mean D-scores and corresponding *p*-values. Recall that positive D-scores indicate pro-stereotypical schemas: relatively strong associations between the Republican/Democratic Party and, respectively, whites/non-whites, religious/non-religious people, rich/poor, and conservatives/liberals. Overall, partisans exhibit the greatest discrimination on racial schemas while agreeing much more strongly about how ideological groups identify with parties. This means that white Democrats often view Democrats as white and therefore push the D score close to 0. In other words, on average, Democrats have no consensual racial stereotype for their party. Republicans, on the other hand, strongly associate their party with whites and the Democratic Party with non-whites. The same is true for religiosity and class, where Democratic D-scores are indistinguishable from 0. Only Republicans have strong group stereotypes about their in-party on these social dimensions, believing their party is religious and wealthy while the Democrats are non-religious and poor.

Ideology is different. Here, both Democratic and Republican partisans hold similar and strong stereotypes about the ideology of both parties: Democrats are strongly linked to liberals and Republicans to conservatives. We suspect this may indicate ideological identity may be the downstream consequence of one or more of the other social group attachments. People across the partisan spectrum understand which social groups are associated with “conservatives” and “liberals.” We suspect these strong mental associations may have formed rather recently, given the results in Study 1, but unfortunately there is no way for us to go back and perform these tests in an era when parties were less sorted ideologically.

[Figure 4 about here]

Next, we explore the association between each of the four group–party schemas (race, religiosity, class, and ideology) and affective polarization. Remember that our causal model presumes the effects of group–party schemas on affective polarization are conditional on the corresponding group attitudes: For instance, the image of the Republican Party as white should translate into pro-Republican affect most powerfully among those with more positive attitudes about whites. For those with negative attitudes toward whites, the same schema would lead to anti-Republican affect. Therefore, we estimate OLS regression models with interactive effects of group–party schemas and group feeling thermometers on partisan affect. In the following analyses, the schema and group attitude variables are normalized to run from zero (the lowest observed value) to one (the highest observed value). The dependent variable, as before, is the Republican feeling thermometer minus Democratic feeling thermometer that can take values from –100 (for respondents who rated the Democratic Party at 100 and the Republican Party at 0), to 100 (for respondents who rated the Democratic Party at 0 and the Republican Party at 100).

Table 4 presents the results of our interactive tests. Note that each respondent could complete only one IAT task and thus had schema scores for only one of the four group dimensions. This means we must estimate four independent OLS models for respondents in each IAT task. We did not include controls in these models since our goal is simply to compare the strength of the association between each party-group schema and affective polarization directly. The results are quite consistent with our theoretical expectations: The effect of the race–party schema is highly moderated by affect toward whites. Recall that positive scores on affective polarization correspond to a preference for the Republican party. In the first row, which corresponds to the effect of the racial schema among those who have very negative attitudes

about whites, perceiving the Republican party to be white and the Democratic party to be non-white lowers one's affect toward Republicans relative to Democrats. In the second row, which is the effect of the white feeling thermometer among those with counter-stereotypic schemas about the racial coalitions of the parties, increasing warmth toward whites reduces positive affect for the Republican party relative to the Democratic party. And in the 3<sup>rd</sup> row, the effect of the white feeling thermometer on affective polarization powerfully switches signs as one's racial schema becomes more strongly stereotypical. That is to say, among those who believe Republicans are white and Democrats are non-white, positive affect toward whites is powerfully associated with affective preference for the Republican party. The interaction is substantially larger than the corresponding effects for social class and ideology—but is nearly identical to the effect of religiosity in the second model (rows 4-6). Comparison tests obtained by running a three-way interaction model (schema–thermometer interaction by group condition) reveal that the interaction between group–party schemas and group attitude for race is significantly greater than that for ideology ( $\Delta\beta = 85.1, p < .05$ ), and class ( $\Delta\beta = 85.6, p < .05$ ), but is statistically indistinguishable from religion ( $\Delta\beta = -18.4, p = .65$ ).

[Table 4 about here]

Since numerical estimates of interactive effects are difficult to interpret substantively, we also present them graphically in Figure 5. The Y axis in each of these figures represents the effect of moving from 0 to 100 on the white feeling thermometer on affective polarization. On the right side of the first panel- capturing people who most strongly associate the Republican Party with whites and the Democratic Party with non-whites- moving from cold to warm on the white feeling thermometer produces a nearly complete positive shift in affect for the Republicans over the Democrats. In other words, respondents who (a) think of Democrats as non-white and

Republicans as white and (b) feel very warmly toward whites as a group effectively rate the Democratic Party at 0 and the Republican Party at 100. For those with the opposite race–party schema (Democrats white, Republicans non-white), the pattern is partially reversed: Moving from low to high on the white thermometer reduces positive affect for the Republicans over the Democrats by about 75 degrees.<sup>5</sup> Note that these respondents are mostly white Democrats in our sample. The corresponding effects are approximately the same size for religion. The differences we see for schemas about social class and ideology are only half as large, and only those with very stereotypical schemas show significant effects of social group affect on partisan polarization.

[Figure 5 about here]

As with Study 2, the results of Study 3 show that race–party schemas are central to the growing affective gap between partisans. In this case, however, all four group–party schemas return significant interactions with group attitudes in predicting affective polarization. The interactive effect of implicit race–party schemas, measured using the original modification of the IAT technique and affect toward whites, is substantively and significantly stronger than the those of social class and ideology. At the same time, the interactive effect of implicit religious-party schemas appears as strong as the race–party linkage. What should we make of this? As in Study 2, we have theoretical and methodological reasons to suspect the impact of the racial schema may be biased downward, relative to other identity dimensions. We discuss these thoughts next.

### **Discussion and Conclusion**

The rise of affective polarization is of significant concern. At the most extreme level, so much antipathy may plausibly increase the chance that partisans will rely on bullets rather than ballots

---

<sup>5</sup> There are very few such individuals in our sample, though being a white Democrat is one very good predictor of holding such a schema.

to settle policy disputes. One could have witnessed an example of this on January 6, 2021, when a violent mob laid siege to the Capitol building, intent on stopping the Congress from ceremonially accepting the results of the presidential election that had taken place months before. Many believe that these events were directly provoked by Donald Trump and other Republicans who questioned the integrity of the 2020 presidential election. Similarly, some of the violence that occurred during the racial justice protests happening over the spring and summer 2020 was defended by politicians and commentators on the left. Evidence from other countries suggests that when politics is seen fundamentally as a competition between ethnic groups, the commitment to winning at any cost, including through the use of violence, grows (Horowitz 1985). If our theory is correct, the imagined ethnic/racial compositions of the two parties have become so distinct that some ardent partisans are willing to call for or rationalize violence in order to maintain political power.

Our theory also has important implications for the conceptualization of partisanship in American politics. Scholars of public opinion have long argued that party identities powerfully structure voting behavior because they are built on fundamental social identities people care about: race, religion, class, and so on (Lazarsfeld, Gaudet, and Berelson 1948). We suspect, however, that race is most central among these divisions in the United States, and that beliefs about “us and them” along racial lines increasingly structure the imagined party coalitions. If we are correct, Republican politicians who reject the “win-at-any-cost” strategy will be considered a traitor to their party because they are challenging the most deeply divisive social group cleavage in American politics.

Across several sources of evidence, we find that both the actual and perceived racial coalitions of the major parties powerfully drive affective polarization. In Study 1, the association

of actual racial partisan sorting with affective polarization from 1952 to 2016 has been extremely pronounced compared to sorting on rivaling social cleavages like religion and objective class. While ideology now rivals race in terms of actual partisan sorting, that sorting seems to have picked up speed in the 1980s, long after the parties began to diverge on race. And the effect of partisan sorting on race tightly predicts affective polarization, again rivaled only by ideology. Comparisons across identity categories made in Study 2 suggest that explicit race–party schemas, compared to ones based on religion or social class, are more important in predicting partisan affective polarization. Once again, only ideological stereotypes are more powerful drivers of affective polarization using explicit measures of party coalitions. In Study 3, implicit schemas linking race and party outperform ones about social class and symbolic ideology, but equal the effect of implicit linkages between religiosity and party. If partisan racial stereotypes are really at the heart of affective polarization, what accounts for these other dimensions sometimes popping in our studies? We think there are a number of reasons.

One major threat to consistently identifying the effect of partisan racial schemas is social desirability bias. Even though the norms of racial rhetoric in American politics seem to be changing (Valentino, Neuner, and Vandebroek 2018), the open expression of racial animosity and stereotyping is still considered socially unacceptable. The latter is not necessarily true for prejudice based on religion, social class, or ideology. Therefore, explicit measures of both race–party schemas and racial attitudes can underestimate the true variances in the sample and the population—meaning the corresponding effects’ estimates in regression analysis may be downward biased.

Another problem we face is the distinct measurement of these group–party schemas in our model. Ideological and religious labels may themselves have become highly racialized in

recent years, clouding our ability to distinguish between them and racial schemas as drivers of affective polarization. We did our best to create tests that distinguish between these group dimensions, but if our respondents thought primarily of white people when they saw the word “Evangelical” in the IAT task, then the tight relationship we found for religiosity may not falsify our hypothesis. The same, of course, may be true for ideology. In other words, racial stereotypes about the parties may be central but hard to measure, and difficult to distinguish empirically from other group dimensions.

Having offered these speculations, however, our evidence still suggests ideological and religious identities may be powerful drivers of affective polarization in the current moment. Class stereotypes about the parties, on the other hand, are consistently weak and fail to cleave the electorate in terms of their affect about the parties. Race, however, is the most consistent of the forces we have examined.

Of course, we have only examined a limited number of the group–party schemas that might have something to do with this increase in hard feelings others have documented. Furthermore, the nature of racial divide itself is changing, as the level of diversity within American society increases. For instance, a shift to the right among white voters may be at least partially related to recent changes in Hispanic/Latino immigration, as opposed to the older conflicts between whites and blacks (Abrajano and Hajnal 2015; Hajnal and Rivera 2014; Ostfeld 2019; but cf. Hopkins 2021). Future work could explore how the parties are imagined in terms of a variety of racial and ethnic groups. In the end, of course, the evidence we have presented is entirely observational, making it impossible to conclusively determine which social group dimensions most powerfully drive affective polarization in a causal sense. The overall

pattern, however, seems to point most heavily at race and perceived racial differences between the two major parties in America.

### References

- Abrajano, Marisa, and Zoltan L. Hajnal. 2015. *White Backlash: Immigration, Race, and American Politics*. Princeton, NJ: 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): 964–81.
- Ansolahehere, Stephen, Jonathan Rodden, and James M. Snyder Jr. 2006. "Purple America." *Journal of Economic Perspectives* 20 (2): 98–118.
- Baldassarri, Delia, and Barum Park. 2020. "Was There a Culture War? Partisan Polarization and Secular Trends in US Public Opinion." *Journal of Politics* 82 (3): 809–27.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20 (3): 351–68.
- Berinsky, Adam J., Justin de Benedictis-Kessner, Megan E. Goldberg, and Michele F. Margolis. 2020. "The Effect of Associative Racial Cues in Elections." *Political Communication* 37 (4): 512–52.
- Bougher, Lori D. 2017. "The Correlates of Discord: Identity, Issue Alignment, and Political Hostility in Polarized America." *Political Behavior* 39 (3): 731–62.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley and Sons.

- Carpenter, Thomas P., Ruth Pogacar, Chris Pullig, Michal Kouril, Stephen Aguilar, Jordan LaBouff, Naomi Isenberg, and Alek Chakroff. 2019. "Survey-Software Implicit Association Tests: A Methodological and Empirical Analysis." *Behavior Research Methods* 51 (5): 2194–08.
- Cassese, Erin C. 2021. "Partisan Dehumanization in American Politics." *Political Behavior* 43: 29–50.
- Claassen, Ryan L., Paul A. Djupe, Andrew R. Lewis, and Jacob R. Neiheisel. 2021. "Which Party Represents My Group? The Group Foundations of Partisan Choice and Polarization." *Political Behavior* 43: 615–36.
- Conover, Pamela Johnston. 1988. "The Role of Social Groups in Political Thinking." *British Journal of Political Science* 18 (1): 51–76.
- Conover, Pamela Johnston, and Stanley Feldman. 1984. "How People Organize the Political World: A Schematic Model." *American Journal of Political Science* 28 (1): 95–126.
- Cramer, Katherine. 2020. "Understanding the Role of Racism in Contemporary US Public Opinion." *Annual Review of Political Science* 23: 153–69.
- Davis, Nicholas T., and Lilliana Mason. 2016. "Sorting and the Split-Ticket: Evidence from Presidential and Subpresidential Elections." *Political Behavior* 38 (2): 337–54.
- Diermeier, Daniel, and Christopher Li. 2019. "Partisan Affect and Elite Polarization." *American Political Science Review* 113 (1): 277–81.
- Filindra, Alexandra, and Noah J. Kaplan. 2016. "Racial Resentment and Whites' Gun Policy Preferences in Contemporary America." *Political Behavior* 38 (2): 255–75.
- Fiorina, Morris P., Samuel J. Abrams, and Jeremy C. Pope. 2006. *Culture War? The Myth of a Polarized America*. New York: Pearson Education.

- Fiske, Susan T., and Patricia W. Linville. 1980. "What Does the Schema Concept Buy Us?" *Personality and Social Psychology Bulletin* 6 (4): 543–57.
- Garrett, Kristin N., and Alexa Bankert. 2020. "The Moral Roots of Partisan Division: How Moral Conviction Heightens Affective Polarization." *British Journal of Political Science* 50 (2): 621–40.
- Gilens, Martin. 1999. *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. Chicago: University of Chicago Press.
- Gillion, Daniel Q., Jonathan M. Ladd, and Marc Meredith. 2020. "Party Polarization, Ideological Sorting and the Emergence of the US Partisan Gender Gap." *British Journal of Political Science* 50 (4): 1217–43.
- Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.
- Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74 (6): 1464–80.
- Greenwald, Antony G., Brian A. Nosek, and Mahzarin R. Banaji. 2003. "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm." *Journal of Personality and Social Psychology* 85 (2): 197–216.
- Hajnal, Zoltan L., and Michael U. Rivera. 2014. "Immigration, Latinos, and White Partisan Politics: The New Democratic Defection." *American Journal of Political Science* 58 (4): 773–89.

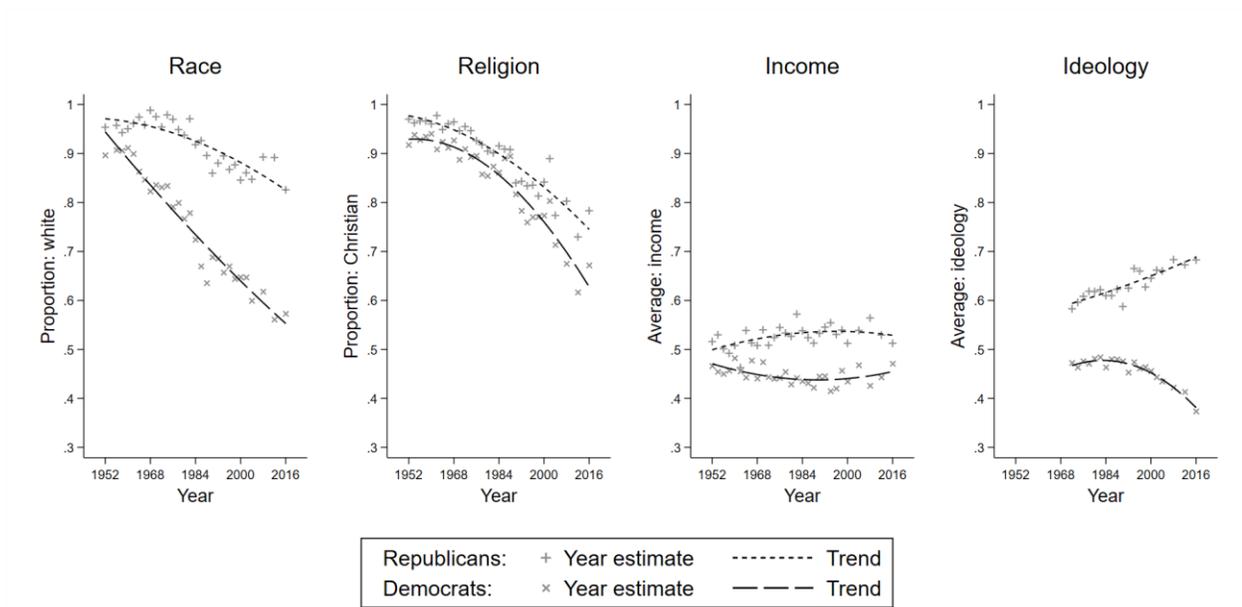
- Hopkins, Daniel J. 2021. "The Activation of Prejudice and Presidential Voting: Panel Evidence from the 2016 U.S. Election." *Political Behavior* 43: 663–86.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Huddy, Leonie, and Omer Yair. 2021. "Reducing Affective Polarization: Warm Group Relations or Policy Compromise?" *Political Psychology* 42 (2): 291–309.
- Hutchings, Vincent L., and Nicholas A. Valentino. 2004. "The Centrality of Race in American Politics." *Annual Review of Political Science* 7: 383–408.
- Iyengar, Shanto, and Masha Krupenkin. 2018. "The Strengthening of Partisan Affect." *Advances in Political Psychology* 39: 201–18.
- 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* 22: 129–46.
- Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76 (3): 405–31.
- 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.
- Kinder, Donald R., and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic Ideals*. Chicago: University of Chicago Press.
- Lazarsfeld, Paul F., Hazel Gaudet, and Bernard Berelson. 1948. *The People's Choice: How the Voter Makes up His Mind in a Presidential Campaign*. 2nd edition. New York: Columbia University Press.

- Lelkes, Yphtach. 2021. "Policy over Party: Comparing the Effects of Candidate Ideology and Party on Affective Polarization." *Political Science Research and Methods* 9 (1): 189–196.
- Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar. 2017. "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect." *American Journal of Political Science* 61 (1): 5–20.
- Lelkes, Yphtach, and Sean J. Westwood. 2017. "The Limits of Partisan Prejudice." *Journal of Politics* 79 (2): 485–501.
- Levendusky, Matthew. 2013. "Partisan Media Exposure and Attitudes Toward the Opposition." *Political Communication* 30 (4): 565–81.
- Lipset, Seymour Martin. 1960. *Political Man: The Social Bases of Politics*. New York: Doubleday and Company.
- Martherus, James L., Andres G. Martinez, Paul K. Piff, and Alexander G. Theodoridis. 2021. "Party Animals? Extreme Partisan Polarization and Dehumanization." *Political Behavior* 43: 517–40.
- 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. 2018. *Uncivil Agreement: How Politics Became Our Identity*. Chicago: 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: 257–77.
- McClain, Paula D. 2021. "Crises, Race, Acknowledgement: The Centrality of Race, Ethnicity, and Politics to the Future of Political Science." *Perspectives on Politics* 19 (1): 7–18.

- Mickey, Robert. 2015. *Paths Out of Dixie: The Democratization of Authoritarian Enclaves in America's Deep South, 1944-1972*. Princeton, NJ: Princeton University Press.
- Nosek, Brian A., Anthony G. Greenwald, and Mahzarin R. Banaji. 2007. "The Implicit Association Test at Age 7: A Methodological and Conceptual Review." In *Automatic Processes in Social Thinking and Behavior*, edited by J. A. Bargh, 265–92. New York: Psychology Press.
- Orr, Lilla V., and Gregory A. Huber. 2019. "The Policy Basis of Measured Partisan Animosity in the United States." *American Journal of Political Science* 64 (3): 569–86.
- Osborne, Danny, David O. Sears, and Nicholas A. Valentino. 2011. "The End of the Solidly Democratic South: The Impressionable Years Hypothesis." *Political Psychology* 31 (1): 81–108.
- Ostfeld, Mara Cecilia. 2019. "The New White Flight? The Effects of Political Appeals to Latinos on White Democrats." *Political Behavior* 41 (3): 561–82.
- Robison, Joshua, and Rachel L. Moskowitz. 2019. "The Group Basis of Partisan Affective Polarization." *Journal of Politics* 81 (3): 1075–79.
- Rogowski, Jon C., and Joseph L. Sutherland. 2016. "How Ideology Fuels Affective Polarization." *Political Behavior* 38 (2): 485–508.
- Rom, Mark Carl, and Paul Musgrave. 2014. "Political Outcome Bias in Grading: Identifying Problems and Suggesting Solutions." *Journal of Political Science Education* 10 (2): 136–54.
- Ryan, Timothy J. 2017. "No Compromise: Political Consequences of Moralized Attitudes." *American Journal of Political Science* 61 (2): 409–23.

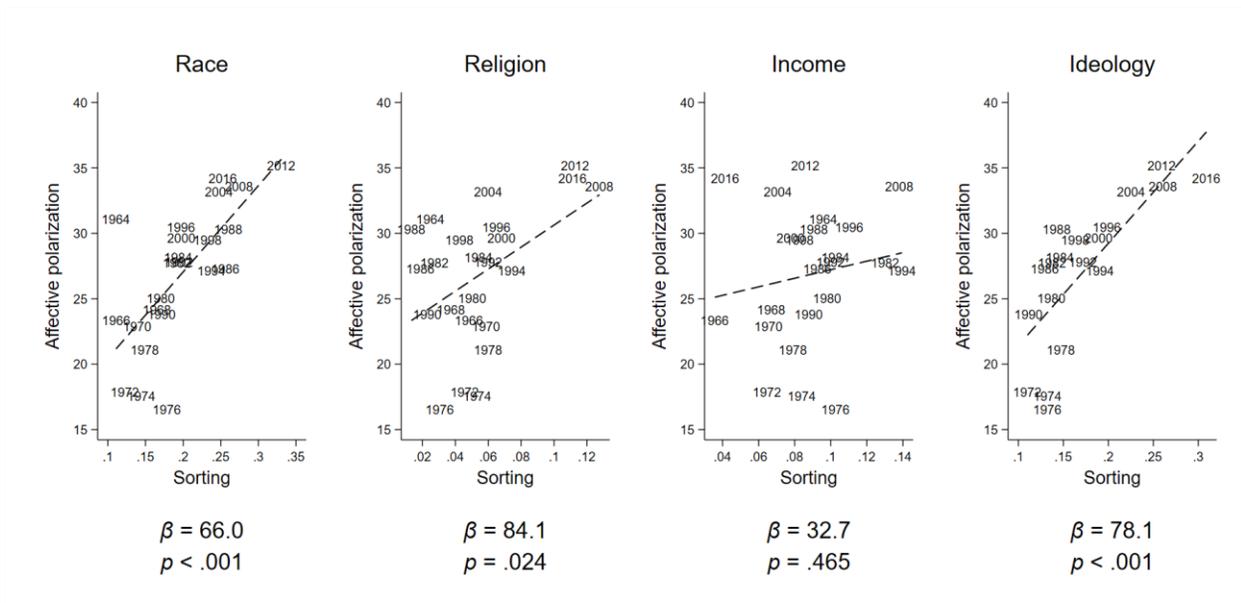
- Shafranek, Richard M. 2020. "Political Consequences of Partisan Prejudice." *Political Psychology* 41 (1): 35–51.
- Strickler, Ryan. 2018. "Deliberate with the Enemy? Polarization, Social Identity, and Attitudes toward Disagreement." *Political Research Quarterly* 71 (1): 3–18.
- Tajfel, Henri, and John C. Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel, 33–47. Monterey, CA: Brooks/Cole.
- Tesler, Michael. 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 (3): 690–704.
- Tesler, Michael. 2016. *Post-Racial or Most-Racial? Race and Politics in the Obama Era*. Chicago: University of Chicago Press.
- Tesler, Michael, and David O. Sears. 2010. *Obama's Race: The 2008 Election and the Dream of a Post-Racial America*. Chicago: University of Chicago Press.
- Valentino, Nicholas A. 1999. "Crime News and the Priming of Racial Attitudes During Evaluations of the President." *Public Opinion Quarterly* 63 (3): 293–320.
- Valentino, Nicholas A., and Ted Brader. 2011. "The Sword's Other Edge: Perceptions of Discrimination and Racial Policy Opinion after Obama." *Public Opinion Quarterly* 75 (2): 201–26.
- Valentino, Nicholas A., Fabian G. Neuner, and L. Matthew Vandebroek. 2018. "The Changing Norms of Racial Political Rhetoric and the End of Racial Priming." *Journal of Politics* 80 (3): 757–71.

- 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): 672–88.
- Voas, David, and Mark Chaves. 2016. "Is the United States a Counterexample to the Secularization Thesis?" *American Journal of Sociology* 121 (5): 1517–56.
- Zhirkov, Kirill, and Nicholas A. Valentino. 2021. "Revisiting the Measurement of Group Schemas in Political Science." In *The Cambridge Handbook of Implicit Bias and Racism*, edited by Jon A. Krosnick, Tobias H. Stark, and Amanda L. Scott. New York: Cambridge University Press.
- Zuckerman, Alan S., Nicholas A. Valentino, and Ezra W. Zuckerman. 1994. "A Structural Theory of Vote Choice: Social and Political Networks and Electoral Flows in Britain and the United States." *Journal of Politics* 56 (4): 1008–33.
- Zuckerman, Alan S, Josip Dasovic, and Jennifer Fitzgerald. 2012. *Partisan Families: The Social Logic of Bounded Partisanship in Germany and Britain*. New York: Cambridge University Press.



**Figure 1.** Compositions of the partisan coalitions on the basis of race, religion, income, and ideology.

*Note:* For race and religion, entries represent the proportion of each party that identifies as white and Christian, respectively. For income, entries represent the mean income of identifiers of each party, where the entire income distribution for each year is recoded to run from 0–1. The midpoint of that scale, .5, therefore represents approximately the middle of the income distribution in a given year. For ideology, the 1–7 scale from extremely liberal to extremely conservative is recoded to run from 0–1, such that .5 represents self-identified moderates.



**Figure 2.** Bivariate relationships between social sorting and affective polarization from 1964 to 2016

*Note:* The X axis plots the differences between party identifiers in terms of race (proportion white), religion (proportion Christian), income (average income), and ideology (average score on the 1–7 liberal conservative scale), all recoded to run from 0–1. The Y axis is the average absolute value of the difference between the two feeling thermometers in each year. Feeling thermometer from 1964 to 1974 are for the Democratic and Republican parties, from 1978 on they are for Democrats and Republicans.

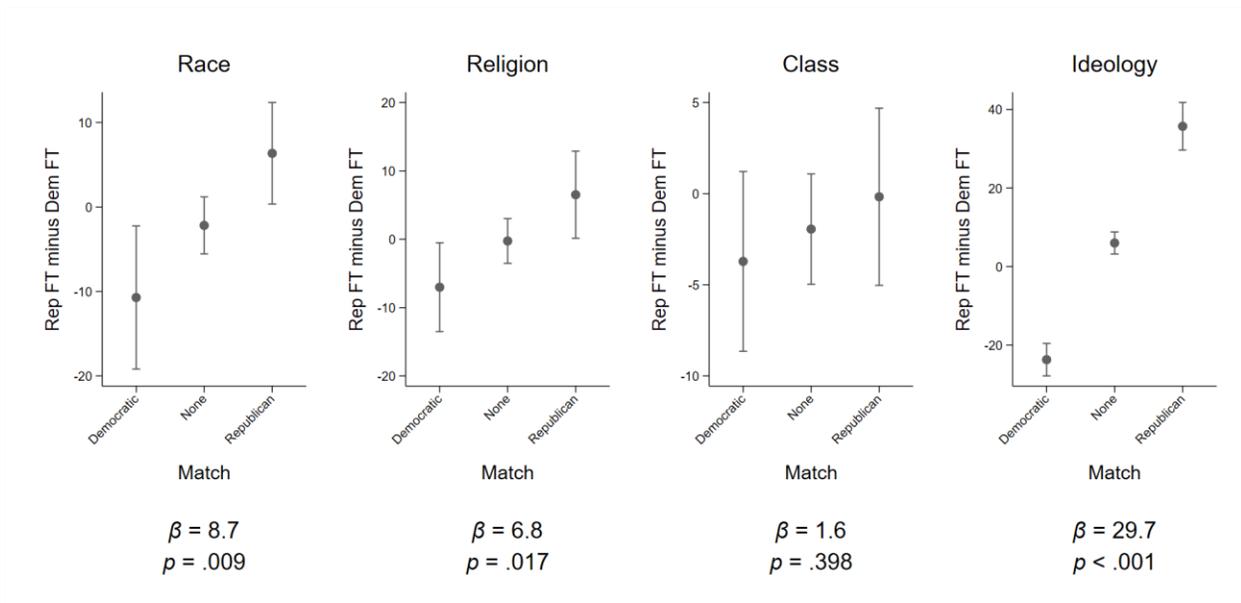
**Table 1.** Frequencies of identity categories associated with “typical” partisans

	Democratic	Republican
Race		
White	65.9	97.4
Black	29.8	0.2
Other	4.3	2.4
Religion		
Christian	55.7	98.0
Secular	38.6	1.5
Other	5.9	0.4
Social class		
Working or lower	57.8	20.9
Middle or upper	42.2	79.1
Ideology		
Conservative	15.9	91.4
Liberal	81.3	2.6
Other	2.8	8.6

*Note.* Cell entries are percentages that total 100% in each row, for each group dimension. Sample sizes for these analyses range from 440 (for the typical race of each party) to 465 (for the typical ideology of each party).

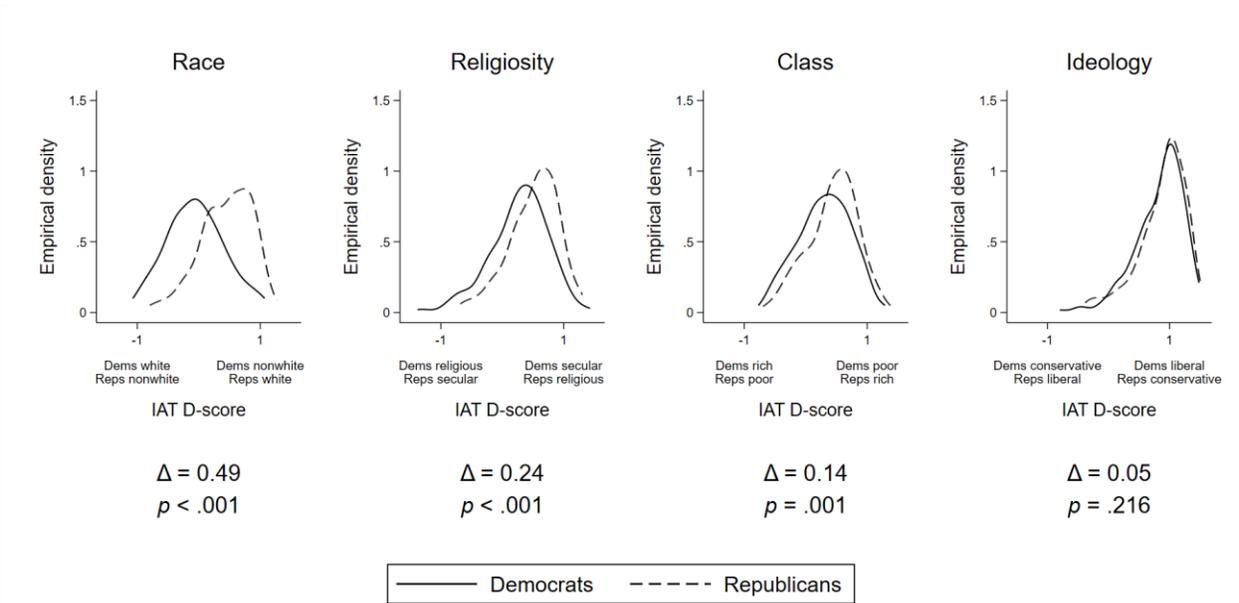
**Table 2.** Construction of the match measures: example

		Respondent is Christian		
		Christian	Republican stereotype All other religions	Secular
Democratic stereotype	Christian	0	-1	-1
	All other religions	1	0	0
	Secular	1	0	0
		Respondent is secular		
		Christian	Republican stereotype All other religions	Secular
Democratic stereotype	Christian	0	0	1
	All other religions	0	0	1
	Secular	-1	-1	0



**Figure 3.** Marginal effects of social group matching on partisan affective polarization across identity and party categories

*Note:* The  $x$  axis in each panel captures whether respondent believes their race (white, black, other), religion (Christian, secular, other), class (working/lower, middle/upper) and ideology (liberal, conservative, moderate) match one of the parties, and whether the outgroups on these dimensions are typical of the other party. Those in “none” believe typical members of each party are the same on a given dimension.



**Figure 4.** Effects of group–party schemas measured as IAT D-scores on affective polarization depending on group affect

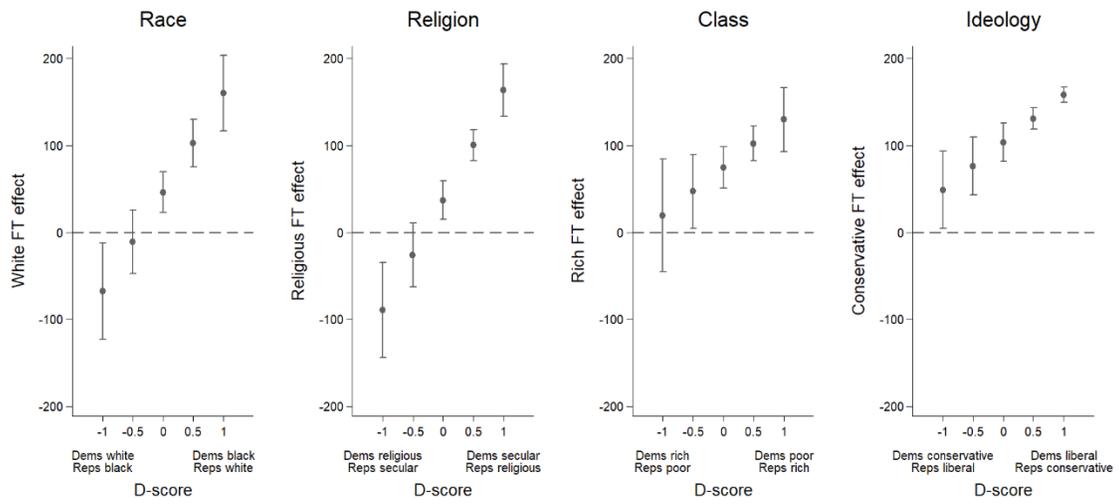
*Note:*  $\Delta$  = difference in schema means by party (Republican mean minus Democratic mean)

**Table 4.** The association of implicit schemas with affective polarization, moderated by group affect

	Estimate
Implicit racial schema	-52.5*
	(23.2)
FT: white people	-55.4***
	(16.8)
Racial schema × White FT	165.1***
	(30.1)
Implicit religious schema	-70.1***
	(18.3)
FT: religious people	-68.5***
	(16.3)
Religious schema × Religious FT	183.5***
	(26.4)
Implicit class schema	-14.7
	(16.8)
FT: rich people	-0.7
	(18.7)
Class schema × Rich FT	80.0*
	(31.1)
Implicit ideological schema	-43.2**
	(13.8)
FT: Conservatives	14.1
	(16.8)
Ideological schema × Conservative FT	79.5**
	(25.4)

*Note.* Standard errors in parentheses. The four models, each represented by a set of three rows in the table above, were run separately.

\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$



**Figure 5.** Effects of group-party schemas measured as IAT D-scores on affective polarization depending on group affect

## Online Appendix

### Study 1: ANES

#### Partisanship

- “Generally speaking, do you usually think of yourself as a Democrat, a Republican, an independent, or what?”
- (If Democrat) “Would you call yourself a strong Democrat or a not very strong Democrat?”
- (If Republican) “Would you call yourself a strong Republican or a not very strong Republican?”
- (if independent) “Do you think of yourself as closer to the Republican Party or to the Democratic Party?”

#### Race

Coded by interviewers.

Three categories in the dataset: *White non-Hispanic*, *Black non-Hispanic*, *Other*.

Recoded to 0 = *Non-white* and 1 = *White* (non-Hispanic).

#### Religiosity

1970–1988: “Would you say you/do you go to (church/synagogue) every week, almost every week, once or twice a month, a few times a year, or never?”

1990 AND LATER: “Lots of things come up that keep people from attending religious services even if they want to. Thinking about your life these days, do you ever attend religious services, apart from occasional weddings, baptisms or funerals? Do you go to religious services every week, almost every week, once or twice a month, a few times a year, or never?”

Answers in the dataset coded from 1 = *Every week* to 5 = *Never*.

Recoded to an interval scale from 0 = *Never* to 1 = *Every week*.

### **Income**

Questions were formulated differently across years. See: <https://electionstudies.org/data-center/anes-time-series-cumulative-data-file/>

Answers in the dataset coded from 1 = *0 to 16 percentile* to 5 = *96 to 100 percentile*.

Recoded to an interval scale from 0 = *0 to 16 percentile* to 1 = *96 to 100 percentile*.

### **Ideology**

“We hear a lot of talk these days about liberals and conservatives. Here is a 7-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale?”

Answers in the dataset coded from 1 = *Extremely liberal* to 7 = *Extremely conservative*.

Recoded to an interval scale from 0 = *Extremely liberal* to 1 = *Extremely conservative*.

### **Feeling thermometers**

“We would like to get your feelings toward some of our political leaders and other people who are in the news these days. We will show the name of a person and we’d like you to rate that person using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don’t feel favorable toward the person and that you don’t care too much for that person. You would rate the person at the 50-degree mark if you don’t feel particularly warm or cold toward the person.”

- Democratic Party
- Republican Party

## Study 2: explicit schemas

### **Respondent's identity: race**

“In terms of race, which of the following comes closest to describing you?”

Answer categories: *White/Caucasian American, Black/African American, Asian American, Native American, Hispanic/Latino American, Other.*

Recoded to *White, Black, Other.*

### **Respondent's identity: religion**

“In terms of religion, which of the following comes closest to describing you?”

Answer categories: *Mainline Protestant, Catholic, Evangelical Christian, Other Christian, Follower of a non-Christian religion, Secular person, Other.*

Recoded to *Evangelical Christian, Secular person, Other.*

### **Respondent's identity: social class**

“In terms of religion, which of the following comes closest to describing you?”

Answer categories: *Middle class, Working class, Other.*

### **Respondent's identity: ideology**

“In terms of religion, which of the following comes closest to describing you?”

Answer categories: *Liberal, Moderate, Conservative, Other.*

Recoded to *Conservative, Liberal, Moderate.*

**Typical partisan: race**

“In terms of race, which of the following comes closest to describing the typical supporter of the [Democratic Party/Republican Party]?”

Answer categories: *White/Caucasian American, Black/African American, Asian American, Native American, Hispanic/Latino American, Other.*

Recoded to *White, Black, “Neutral.”*

**Typical partisan: religion**

“In terms of religion, which of the following comes closest to describing the typical supporter of the [Democratic Party/Republican Party]?”

Answer categories: *Mainline Protestant, Catholic, Evangelical Christian, Other Christian, Follower of a non-Christian religion, Secular person, Other.*

Recoded to *Evangelical Christian, Secular person, “Neutral.”*

**Typical partisan: social class**

“In terms of social class, which of the following comes closest to describing the typical supporter of the [Democratic Party/Republican Party]?”

Answer categories: *Middle class, Working class, Other.*

Recoded to *Middle class, Working class.*

**Typical partisan: ideology**

“In terms of ideology, which of the following comes closest to describing the typical supporter of the [Democratic Party/Republican Party]?”

Answer categories: *Liberal, Moderate, Conservative, Other.*

Recoded to *Conservative, Liberal, Moderate.*

**Match measure: race**

		Respondent is white		
		White	Republican stereotype "Neutral"	Black
Democratic stereotype	White	0	-1	-1
	"Neutral"	1	0	0
	Black	1	0	0
		Respondent is black		
		White	Republican stereotype "Neutral"	Black
Democratic stereotype	White	0	0	1
	"Neutral"	0	0	1
	Black	-1	-1	0

The "Neutral" category includes all racial/ethnic categories other than black and white.

**Match measure: religion**

		Respondent is Evangelical		
		Evangelical	Republican stereotype "Neutral"	Secular
Democratic stereotype	Evangelical	0	-1	-1
	"Neutral"	1	0	0
	Secular	1	0	0
		Respondent is secular		
		Evangelical	Republican stereotype "Neutral"	Secular
Democratic stereotype	Evangelical	0	0	1
	"Neutral"	0	0	1
	Secular	-1	-1	0

The "Neutral" category includes all religious affiliations other than Evangelical Christian and secular

**Match measure: social class**

		Respondent is middle-class	
		Republican stereotype	
		Middle class	Working class
Democratic stereotype	Middle class	0	-1
	Working class	1	0

		Respondent is working-class	
		Republican stereotype	
		Middle class	Working class
Democratic stereotype	Middle class	0	1
	Working class	-1	0

There was no “Neutral” category for social class

**Match measure: ideology**

		Respondent is conservative		
		Republican stereotype		
		Conservative	Moderate	Liberal
Democratic stereotype	Conservative	0	-1	-1
	Moderate	1	0	0
	Liberal	1	0	0

		Respondent is liberal		
		Republican stereotype		
		Conservative	Moderate	Liberal
Democratic stereotype	Conservative	0	0	1
	Moderate	0	0	1
	Liberal	-1	-1	0

**Size of government**

“Some people think the government should provide fewer services even in areas such as health and education in order to reduce spending. Suppose these people are at one end of a scale, at point 1. Other people feel it is important for the government to provide many more services even if it means an increase in spending. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6. Where would you place yourself on this scale?”

### **Defense spending**

“Some people believe that we should spend much less money for defense. Suppose these people are at one end of a scale, at point 1. Others feel that defense spending should be greatly increased. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6. Where would you place yourself on this scale?”

### **Environmentalism**

“Some people think the federal government needs to regulate business to protect the environment. They think that efforts to protect the environment will also create jobs. Let us say this is point 1 on a 1-7 scale. Others think that the federal government should not regulate business to protect the environment. They think this regulation will not do much to help the environment and will cost us jobs. Let us say this is point 7 on a 1-7 scale. And of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6. Where would you place yourself on this scale?”

### **Abortion**

“Some people think abortion should never be permitted. Let us say this is point 1 on a 1-7 scale. Others think that a woman should always be able to obtain an abortion. Let us say this is point 7 on a 1-7 scale. And of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6. Where would you place yourself on this scale?”

### **Feeling thermometers**

Same as in Study 1.

### Study 3: implicit schemas

#### **Partisanship**

Same as in Study 1.

#### **IAT stimuli: Democratic Party**

Available upon request.

#### **IAT stimuli: race**

*White people:* White, European American, Caucasian, Anglo

*Non-white people:* Black, African American, Minority, Latino

#### **IAT stimuli: religion**

*Religious people:* Evangelical, Religious, Churchgoer, Worshiper

*Non-religious people:* Secular, Non-religious, Atheist, Non-believer

#### **IAT stimuli: social class**

*Rich people:* Rich, Wealthy, Well-off, Upper class

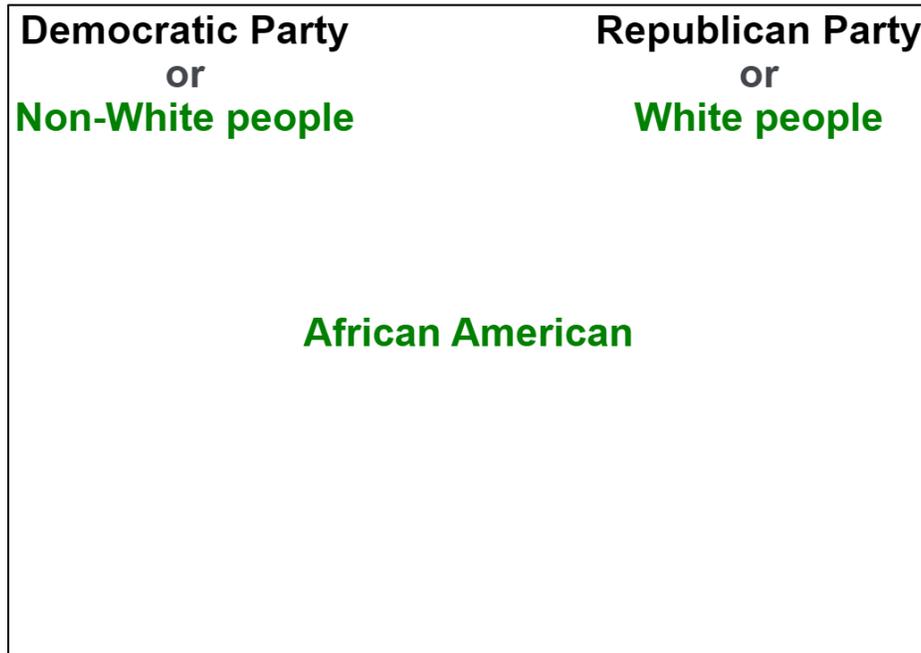
*Poor people:* Poor, Needy, Broke, Lower class

#### **IAT stimuli: ideology**

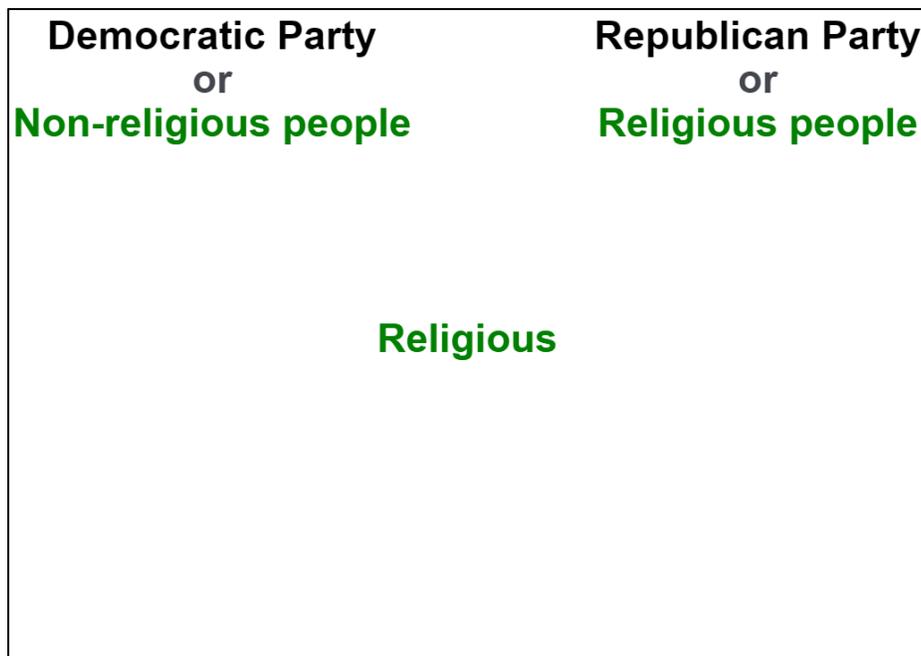
*Conservatives:* Conservative, Right-wing, Rightist, Capitalist

*Liberals:* Liberal, Left-wing, Leftist, Socialist

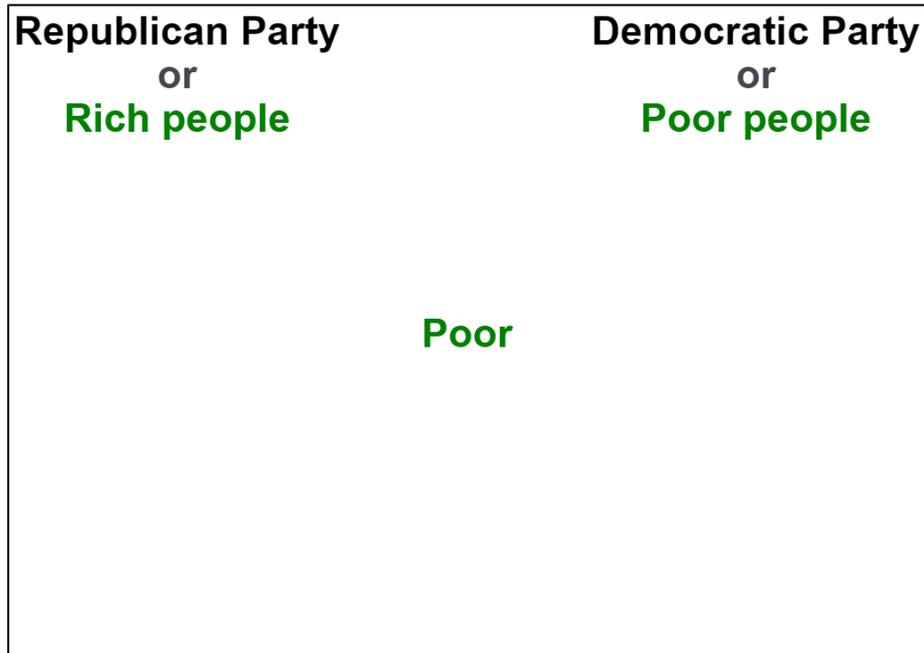
Sample IAT screen: race



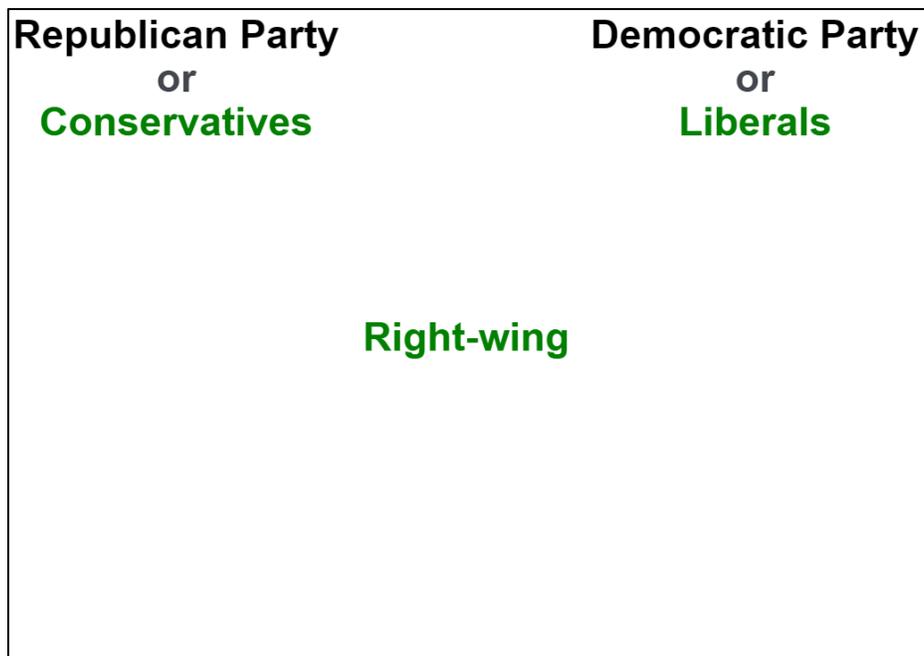
Sample IAT screen: religion



**Sample IAT screen: social class**



**Sample IAT screen: ideology**



**Feeling thermometers**

Same as in Studies 1 and 2.