Americans, Not Partisans: Can Priming American National Identity Reduce Affective Polarization?

Running Head:
Americans, Not Partisans

Author:
Matthew S. Levendusky

Contact Information:
208 S. 37th St.
Philadelphia, PA 19104
(215) 898-7653
mleven@sas.upenn.edu
Abstract:

In recent years, Americans have become more affectively polarized: that is, ordinary Democrats and Republicans increasingly dislike and distrust members of the opposing party. Such polarization is normatively troubling, as it exacerbates gridlock and dissensus in Washington. Given these negative consequences, I investigate whether it is possible to ameliorate this partisan discord. Building on the Common Ingroup Identity Model from social psychology, I show that by heightening subjects’ sense of American national identity, they come to see members of the opposing party as fellow Americans, rather than rival partisans. As a result, they like the opposing party more, thereby reducing affective polarization. Using several original experiments, as well as a natural experiment surrounding the July 4th holiday and the 2008 Summer Olympics, I find strong support for my argument. I conclude by discussing the implications of these findings for efforts to reduce polarization more generally.

Keywords:
American identity, Common ingroup identity model, affective polarization, partisanship

Note:

Supplemental material for this article is available in the appendix in the online edition. Replication files are available in the JOP Data Archive on Dataverse (http://thedata.harvard.edu/dvn/dv/jop). The experimental study described herein was reviewed by the University of Pennsylvania Institutional Review Board and deemed to be exempt.
Americans, we are told, are a polarized people. While initially only political elites were thought to be polarized (McCarty, Poole, and Rosenthal 2006), now, by some measures, so are ordinary Americans. In particular, there is strong evidence that Americans are affectively polarized—Americans now dislike and distrust members of the opposing party (Iyengar, Sood, and Lelkes 2012, Abramowitz and Webster 2016). This affective polarization makes governance more difficult. It leads individuals to distrust government, especially when the opposing party is in control, thereby exacerbating polarization and gridlock in Washington, D.C. (Hetherington and Rudolph 2015). Affective polarization is not the sole—or even primary—cause of government dysfunction, but it certainly contributes to it.

This prompts an important question: is there any remedy to such affective polarization? Drawing on work from social psychology, I explain how subtle psychological primes can reduce affective polarization. In particular, I show that by heightening respondents’ sense of American national identity, they come to see those from the other party as fellow Americans more than members of an opposing political tribe. As a result of this heightened national identity, primed respondents view members of the opposing party more positively, thereby reducing affective polarization. I demonstrate this using original experiments, as well as with data from a natural experiment exploiting the exogenous increase to national identity caused by the July 4th holiday. In both the experimental and real-world data, I find strong support for my argument that increasing American identity decreases affective polarization in the mass public.
Such findings have important implications for what they can tell us about the ability to mitigate polarization in the mass public, and what strategies might be effective in that effort. While this strategy—or any other—is not a panacea, it offers a mechanism that reduces divisions among ordinary citizens.

**American Identity as a Mechanism to Reduce Affective Polarization**

Over the past 15 years, scholars have extensively debated whether the electorate has become more polarized along ideological lines, with some arguing that it has become more divided (Abramowitz 2010), while others claim it has not (Fiorina, Abrams, and Pope 2005). While scholars do not agree on the level of ideological polarization, there is broad consensus that the mass public has become more affectively polarized in recent years. Affective polarization refers to a tendency to dislike and distrust the opposition, and to impute negative characteristics to them (Iyengar, Sood, and Lelkes 2012, Mason 2013). So even if Americans are not deeply divided ideologically, the mass parties have growing antipathy toward one another. For example, the average rating given to members of the opposing party on the 100-point feeling thermometer scale used by the National Election Study dropped 15 degrees between 1988 and 2008 (Iyengar, Sood, and Lelkes 2012, 412-3; see also Abramowitz and Webster 2016). Since 2008, such animosity has only increased—for the first time ever, a majority of partisans now view the other party very unfavorably (Pew 2016), and according to one study, nearly one-half of Americans got into an argument about the 2016 election with a friend, family member, or coworker (Edwards-Levy 2016). By some measures, partisan affective polarization now is more pronounced than racial animus (Iyengar and Westwood 2015). Clearly, much as there is
hostility between the elite parties in Washington, there is also hostility between ordinary Americans as well.

Such discord and affective polarization is normatively troubling. A spirit of compromise—where citizens respect one another’s ideas and agree to find common ground to solve problems—is necessary for a pluralistic polity such as the U.S. to function effectively (Gutmann and Thompson 2012). Affective polarization—with heightened distrust and dislike of the opposition—threatens this sort of decision-making, with consequences that shape the quality and quantity of public policies as well. Affective polarization lowers trust in government, especially when the opposing party is in charge—if you do not like the other side, why would you trust them? When the out-party does not trust the government, they do not encourage their leaders to compromise with the party in power, and gridlock results (Hetherington and Rudolph 2015, Hetherington 2015). While obviously no one would attribute all of the dysfunction in the nation’s capital to affective polarization, it certainly is a factor contributing to “why Washington won’t work,” in Hetherington and Rudolph’s (2015) apt phrasing. Affective polarization has real, negative consequences for governing.

This affective polarization stems from individuals’ identification with a political party. In politics, we have many different identities, and they vary across situations (Huddy 2001). But, at least in the American context, arguably no identity is more important than one’s partisan identity: because elites structure politics in partisan terms, partisan identity sits at the core of political identity (Green, Palmquist, and Schickler 2002). As a result of identifying with a political party (or any other social group), we divide the world into in-groups (our own party) and out-groups (the other party; see
Tajfel and Turner 1979, Tajfel 1981). Lumping the world into in-groups and out-groups gives rise to in-group favoritism and bias (*ibid.*), the tendency to view in-group members (one’s own party) positively, and out-group members (the opposing party) negatively. The seeds of affective polarization are thus sown into the very nature of how we identify as partisans.

The nature of the contemporary political environment magnifies this general tendency. As elites have polarized, the messages they send to voters have clarified, and partisanship has resurged in the mass electorate (Hetherington 2001, Levendusky 2009). Further, because the two parties are relatively balanced in the nation (though not in every region), there is intense competition in every election for control of government. Thus, contemporary politics both makes partisan identities more salient, and it makes it clear that the parties are in competition for resources—the control of government, and with it, the ability to set policy. Both of these factors exacerbate our inherent tendency toward in-group favoritism, and hence affective polarization (Mullen, Brown, and Smith 1992, Brewer 1991). Given this political environment, it is not terribly surprising that affective polarization has increased so dramatically in recent years.

This identity-based understanding of affective polarization also suggests a mechanism for mitigating it. Affective polarization stems from the strength and centrality of partisanship to American’s political identities. But partisanship is not the only political identity Americans possess—indeed, we all have multiple identities that vary in strength, and emphasizing one can reduce the salience and effects of another (Gaertner and Dovidio 2000). If I can therefore prime another political identity, especially one shared by both Democrats and Republicans, I may be able to reduce the salience of partisanship,
and thereby lessen affective polarization and discord. A shift in the political relevant identities may ameliorate polarization.

Given the centrality of partisanship, this task is easier said than done. Precisely because partisanship is so central to our political identities, few other identities can trump it, and that list narrows even further once you restrict it to identities shared by members of both major political parties. But there is one identity than is both powerful and broadly shared by all Americans: their American national identity. While there are many dimensions to American identity (see, among many others, Smith 1997, Schildkraut 2011), here, by American national identity, I mean “a subjective or internalized sense of belonging to the [American] nation” (Huddy and Khatib 2007, 65). This is not tied, however, to an endorsement of a particular ideology or political belief, but rather reflects “being or feeling American” (Huddy and Khatib 2007, 65; see also Theiss-Morse 2009). Ordinary citizens think that this sense of belonging is an important part of being “American.” Schildkraut (2011) asks respondents which factors they think should be important to being an American. More than 90 percent say that “feeling American” and “thinking of oneself as American” should be important components of being an American (Table 3.1, p.45). This sense of national identity transcends racial, ethnic, and partisan/ideological boundaries (Citrin, Duff, and Wong 2001, Huddy and Khatib 2007, Theiss-Morse 2009), and reflects a broadly held conception of being American. Further, American identity is also a strong identity, one capable of rivaling partisanship: 50% of Americans strongly identify as Americans, a significantly larger percentage than those who strongly identify with their religion, race, or sex (Theiss-Morse 2009, figure 2.1,
Clearly, most Americans see their national identification as a core component of their political identity.

The Common Ingroup Identity Model explains how identifying as an American can lessen the influence of partisanship and therefore ameliorate affective polarization (Gaertner et al. 1989, Gaertner and Dovidio 2000). This model argues that individuals have multiple identities, which vary in strength and relevance depending upon the situation. The identity that is most salient will be the one that shapes how people perceive the political world. Normally, when a Democrat thinks about Republicans, her partisan identity is her most salient identity, stemming from its centrality to political thinking in the American context. As a result, she sees Republicans as members of a disliked out-group, and evaluates them negatively. But suppose that I primed that Democrat’s American identity. Now her American identity is more salient than her partisan one, and she therefore perceives Republicans differently. She re-categorizes them, moving them from disliked out-group to liked in-group—they are fellow Americans, not members of an opposing political tribe. This re-categorization triggers her to feel more warmly towards them and see them as more similar to herself (Tajfel and Wilkes 1963). This more positive re-categorization, then, lessens her antipathy toward the other side, and hence vitiates affective polarization. So by changing the salience of an individual’s identities—partisan versus American—one can change how they evaluate those from the other party.

1 This is not the first application of this model to political science: Kam and Ramos (2008) use it to study presidential approval, and Transue (2007) uses it to study racial tolerance. Reid (2012) uses a similar argument (though a slightly different theoretical model) to show effects on hostile media bias. Here, I extend these arguments by using it to study affective polarization.
This effect should extend not only to evaluations of the opposing party, but also to its leaders as well. So, for example, if asked to evaluate President Obama, Republicans who have had their American identity primed should rate the president more warmly, but this sort of national identity prime should have no effect on Democrats.

This same logic implies that the treatment should have no effect on same-party affective evaluations. Members of one’s own party are always part of the in-group: absent any national identity priming, they are part of the partisan in-group, and with national identity priming, they are part of the national in-group. In short, same-party individuals are not re-categorized as a result of the treatment, so there should be no effect in how individuals perceive them.

This leads, then, to a straightforward empirical hypothesis: priming American national identity will reduce affective polarization by triggering a psychological re-categorization of the opposing party (hypothesis 1). More specifically, priming national identity will improve evaluations the opposing party (hypothesis 1A), have no effect on evaluations of the same party (hypothesis 1B), and these same effects will extend to evaluations of political leaders (hypothesis 1C).

But the effect will not be equally strong for all respondents. Such effects should be more muted among those who are sorted (i.e., those whose overall ideological outlook matches their partisan orientation, see Levendusky 2009). As Mason (2015) illustrates, those who are sorted (in her words, have aligned identities) show more hostility toward the other party and see them as more of a threat. Given that party and ideology align and are quite strong in this group, their identity as a partisan is a strong and ingrained one. Because their partisan identities are so entrenched, prioritizing their American identity
will be more difficult (since they will want to hold on to that partisan/ideological way of thinking). Given this, such sorted partisans will be more resistant to this type of priming (hypothesis 2).

The same logic should apply equally well to those who identify most strongly with a party—that is, strong partisans. Here, their stronger partisan identity will serve as a similar anchor, making it more challenging to substitute the American identity for the partisan one. This type of prime should therefore also have a smaller effect on strong partisans (hypothesis 3).

**Experimental Design and Measures**

I test these hypotheses using an original survey experiment. In the pre-test portion of the study, subjects are asked to state their partisan identity and some background demographic attributes (see the appendix for the full questionnaire). The treatment stimulus asks respondents to first read an article about the strengths of America and Americans, modeled on an actual news article (Novarro 2014). Subjects then write a brief paragraph explaining what people like best about America and why they are proud to identify as American. This gives a two-part prime: all subjects read the text-based news article, and its point is then reinforced in the open-ended prime. This ensures that subjects’ identity as an American has been effectively primed. Subjects in the control condition are asked to read a brief apolitical news story and write a short paragraph about an apolitical topic (to effectively mimic the style and length of the treatment).

---

2 This study was pre-registered with EGAP as study 20150713AA; see the online appendix for the deviation from the pre-analysis plan.
There is an important limitation to this treatment, however. Because the news article emphasizes positive dimensions of American life, it almost certainly also primes patriotism—“a deeply felt affective attachment to the nation” (Conover and Feldman 1987, 1). While patriotism and American identity are distinct concepts theoretically (Huddy and Khatib 2007, Theiss-Morse 2009), empirically, they are deeply intertwined. Americans see patriotism as a group norm—Americans are proud of their country. Unsurprisingly, the two are highly correlated: Huddy and Khatib (2007) report a pairwise correlation of 0.74, suggesting a strong empirical correspondence (Table 1, p. 68). Further, those who identify more strongly as Americans are more patriotic (Theiss-Morse 2009, 52-55). So here, by priming American identity, I likely also prime patriotism as well because they are so connected. While this prime—or any other like it—cannot fully disentangle them, the evidence I discuss below suggests that American identity is at least an important component of these effects.

To ensure that the treatment prime actually increased American national identity, I included a series of items measuring American national identity in the post-test questionnaire. The items measure the extent to which respondents perceives themselves to be Americans (i.e., do they use “we” rather than “they” when talking about Americans; the items come from Huddy and Khatib 2007).³

To measure affective polarization, I rely on three main measures. First, I ask respondents to rate both the Democratic and Republican Parties, as well as President

³ Because simply asking people about their American identity might prime American identity (see experiment 3 below), the manipulation check items came at the close of the study after I measure my dependent variables.
Obama, on the standard 100-point feeling thermometer. Second, I ask subjects to rate how well 9 different traits describe the opposing party: American, intelligent, honest, open-minded, generous, hypocritical, selfish, and mean (the battery comes from Garrett et al. 2014). Third, I ask subjects to list their likes and dislikes of the opposing party. All three measures have been used by previous studies of affective polarization: Iyengar et al. (2012) use feeling thermometers and a trait battery, Levendusky and Malhotra (2016) use party likes and dislikes and feeling thermometers, and Garrett et al. (2014) use a trait battery.

If my core hypothesis (hypothesis 1) is correct, then I would expect that after being exposed to the treatment prime, subjects will show lower levels of affective polarization: they will have higher feeling thermometer ratings of the opposing party, evaluate the opposing party more positively along a variety of different dimensions, and have more likes and fewer dislikes of the opposing party. In the analyses, except where the hypothesis predicts a differential effect by party (e.g., H1C about evaluations of President Obama), I pool Democrats and Republicans in the analysis below. I have re-analyzed the data looking for differential effects by party, and I do not find them (see the online appendix), so I pool the parties in the interest of simplicity.

I conducted the experiment on 1,729 respondents interviewed by Survey Sampling Incorporated (SSI) between 15 - 18 July 2015. SSI is an Internet survey firm that maintains a large opt-in panel of respondents who complete surveys in exchange for

---

4 Here, I count Independent partisan leaners as partisans (Keith et al. 1992), though excluding them does not change my substantive results (see the online appendix). I exclude the small percentage of respondents who are pure Independents as I lack a theoretical prediction for them.
compensation. While the sample is not a random sample of Americans, the sample is broadly demographically similar to the overall U.S. population. Any data that is not from a random (probability) sample raises questions about the generalizability of the effects. Here, such concerns are likely to be limited because American identity does not differ greatly by demographics, and there is no ex ante reason to expect differential responsiveness based on observable demographics (and indeed, looking at the data, I do not find such patterns). A lack of a probability sample is therefore unlikely to change my substantive conclusions very much. Further, SSI data compares favorably to other similar firms, and has been used widely throughout political science (for example, Berinsky, Margolis, and Sances 2014).

**Experimental Results**

Before examining the effects of the prime on out-party affect, I need to verify that the treatment prime actually increased American identity. Using the manipulation check items, I find that it does: American identity increases from 4.3 in the control condition to 4.4 in the treatment condition (on a 1-5 scale), a statistically significant increase (t=2.8, p < 0.01). While the effect is relatively modest, this is due to a ceiling effect, given that even in the control condition, subjects already strongly identified as Americans.

Given that the manipulation succeeded, the core question is whether the treatment actually mitigates out-party animus, as I predicted in hypothesis 1. As I explained above,

---

5 The sample is 57% Democratic (including leaners), 52% female, 75% Caucasian, 13% African-American, 16% Hispanic, and the median age is 44 (with one-quarter of the sample being 31 or younger, and one-quarter being 58 or older).
individuals should more positively evaluate the opposing party and its leaders, but there should be no effect on same-party or same-party leader evaluations. Here, that implies that treated Republicans (Democrats) should rate Democrats (Republicans) more warmly on the feeling thermometer scale, while there should be no change to same-party feeling thermometer ratings. Further, treated Republicans should rate President Obama more warmly on the feeling thermometer scale, while there should be no effect for Democrats. Columns 1 through 4 of Table 1 presents the results.

[Insert Table 1 about here]

Overall, the results show strong support for my underlying hypothesis. Increasing American identity makes individuals more positive toward the out-party by approximately 5 degrees. To put this effect in perspective, in the control condition, strong partisans are about 9 degrees cooler toward the out-party than weak or leaning partisans. So the treatment effect is ½ the size of the gap between strong and weak/leaning partisans, which is not a small effect. This effect also extends to the party’s leaders as well—Republicans become 4 degrees warmer toward President Obama. By re-categorizing the out-party as part of a broader in-group (Americans), rather than a partisan out-group, the prime helps to moderate animus toward the out-party.

But given the nature of the feeling thermometer scale, it can be difficult to understand the magnitude of a 4-5 degree feeling thermometer shift. To put this into context, consider what fraction of respondents rate the opposing party at 0 degrees—the least favorable rating. In the control condition, 13 percent of subjects assign the other party this rating, but that falls to 9.5 percent in the treatment condition, a relative decrease of more than 25 percent. Alternatively, consider the percentage of respondents who rate
the opposing party 50 degrees or higher (that is, neutral or positive). Here, that percentage jumps from 17 percent among those in the control group to 23 percent among those in the treatment group, a relative increase of 35 percent. Finally, consider the difference between same-party and out-party feeling thermometer ratings as another manifestation of affective polarization. Here, that gap is 40 degrees in the control condition, but only 35 degrees in the treatment condition, a decline of just over 10 percent (and the partisan gap in evaluations of President Obama shifts by a similar amount). These effects underline the substantive point made above: this is a meaningful shift in how viewers perceive the opposing party.

While there is a large effect on opposite-party feeling thermometers, there is no effect on same-party evaluations: there is a null effect on same-party feeling thermometers, and Democratic evaluations toward President Obama become no warmer as well. Because same-party individuals were always part of the (liked) in-group, the treatment does not change attitudes toward them. This also helps to rule out an alternative explanation that the prime simply made subjects more positive toward all groups.

The results above show that the prime shapes feeling thermometer ratings, and by the same logic, it should also affect trait evaluations. If my argument is correct, treated Democrats (Republicans) should think that Republicans (Democrats) are more American, intelligent, honest, open-minded, and generous, and less hypocritical, selfish, and mean. I

6 Another way to see the relative effect size here is to consider it relative to a factor that we know increases affective polarization. Lelkes, Sood, and Iyengar (2017) show that increasing political interest (a strong driver of affective polarization) increases it 14%. Here, that suggests that my treatment effects are approximately one-third that size, which again is not a small effect.
begin by considering how the treatment affects the overall set of items; column 5 of Table 1 presents the results averaging across all 9 traits ($\alpha=0.80$). Treated subjects rate the other party more positively: substantively, this increase is approximately $1/5^{th}$ of a standard deviation (or slightly less than $1/5^{th}$ of a scale point on the 1-5 scale). But more interestingly, I can also examine how the treatment shaped evaluations of the out-party on the positive traits (American, intelligent, open-minded, honest, and generous), and the negative traits (hypocritical, selfish, and mean); I do so in columns 6 and 7 of Table 1. Here, I find a particularly interesting pattern: the treatment condition increases positive ratings, but it does not decrease negative ratings. That is, treated subjects think the out-party is more American, intelligent, honest, open-minded, and generous (the positive traits), but they do not think those in the out-party are less hypocritical, selfish, or mean. I lack a clear theoretical expectation as to why this is the case, but it presents an interesting possibility of differential effects for future work.

There is one trait, however, that merits additional attention. I asked respondents to assess how “American” the other party is, which also functions as a further manipulation check of my theory. The core causal logic of my theory is that the treatment prime causes subjects to re-categorize members of the out-party as members of a liked in-group (Americans), rather than members of a disliked out-group (out-party partisans). If this is the case, then subjects in the treatment condition should think the label “American” more accurately describes out-party partisans. This is exactly what I find in column 8 of Table 1: treatment subjects think American is a significantly better descriptor of the out-party ($t=3.7$, $p < 0.01$; the effect size is similar to the effects on positive traits overall). The prime does, in fact, cause individuals to see the out-party as part of a liked national in-
group, rather than disliked out-party group. By priming national identity, subjects positively re-categorize out-party partisans as fellow Americans.

As a final measure of affective polarization, I turn to the likes/dislikes items; the results are presented in columns 9 and 10 of Table 1. Subjects offer more likes about the other party, but they offer no fewer dislikes. This directly parallels the findings on trait evaluations above: subjects have more positive things to say, but not necessarily fewer negative things. Again, I leave this discrepancy for future work.

Overall, however, the pattern of results here is quite strong across the different measures, and a clear picture emerges: the prime leads subjects to view the out-party as part of a liked national in-group, rather than seeing them as a disliked partisan out-group. As a result, subjects evaluate them more positively, and have lower levels of animus toward them: they rate them more positively on feeling thermometers, evaluate them more positively on various traits, and have more things that they like about them.7

Hypotheses 2 and 3, however, predict that there will be weaker treatment effects for those who have stronger, more established partisan identities: those who are strong partisans or are sorted (i.e., have an ideology that aligns with their partisan identification). Because these individuals have stronger partisan identities, and hence view the other party more negatively, they should be more resistant to this sort of identity-based priming. Table 2 presents a simplified version of Table 1, testing for these differential treatment effects.

7 One interesting difference is that the effects on feeling thermometers are much larger than the effects on the trait ratings and like/dislike items. The reason why is unclear, but that is an important topic for future studies.
Fascinatingly, I find no differential treatment effects in any case for either strong partisans or for the sorted. Across all three measures (feeling thermometers, trait evaluations, and likes/dislikes), I find no heterogeneous treatment effects. While there is a negative main effect of sorting/strong partisans in every case (i.e., in the control condition, these individuals evaluate the other party more negatively, consistent with previous work), there is no difference in the treatment effects. This lack of effect is not simply due to a lack of statistical power to detect interactive effects, as nearly all of the interaction terms are quite small and are relatively precisely estimated (with the possible exception of the feeling thermometer data for strong partisans). The data do not support the heterogeneous treatment effects predicted by H2 and H3. Even those who are most susceptible to polarized affective evaluations are affected by this prime.  

Additional Experimental Evidence

The results from the experiment above (hereafter, experiment 1) demonstrate that priming American identity (via reading an article about America) decrease animus toward the other party. But it comes with an important limitation as well: the newspaper article is multi-faceted, and may well inadvertently trigger attitudes beyond simply American identity.

---

8 An interesting extension for future work will be to explore additional heterogeneity along these lines. For example, it may be that those who are not just sorted by have a full set of aligned issue preferences may not be susceptible to these effects. I thank an anonymous referee for this point.
To overcome this limitation, I fielded two additional experiments. First, I fielded an experiment where I only used the text-based prime as the treatment: that is, I asked subjects why they are proud to be Americans, without having them first read the newspaper article (I refer to this experiment as experiment 2). Allowing subjects to explain why they prize a particular value is a common technique in social psychology for priming these sorts of constructs (McQueen and Klein 2006). But even this strategy has a limitation, as asking subjects why they are proud of being American may also prime positive dimensions of that identity—the treatment may enhance American identity, rather than simply increasing it. To circumvent this problem, I ran an additional experiment where the treatment simply asked subjects to complete the manipulation check items from experiment 1 above, which measure how strongly they identify as Americans (I refer to this experiment as experiment 3). This is perhaps the simplest and most direct way possible to simply prime American identity. While no treatment is ever perfect, showing consistent results across these different treatments provides a nice robustness check on my results.

I conducted both follow-up studies in Amazon’s Mechanical Turk, a frequent source for experimental subjects that yields results similar to those from other subject pools (Berinsky, Huber, and Lenz 2012, Clifford, Jewell, and Wagoner 2015); see the appendix for additional details and full question wording.9

9 Technically, I ran three new studies: 2 versions of experiment 2, and 1 version of the experiment 3. I pool responses from both versions of experiment 2 because they use identical wordings, and analyzing them separately yields equivalent results, just with lower statistical power. Experiment 2A (N=431) was conducted 16 December 2014; Experiment 2B (N=341) was conducted on 21 March 2016; Experiment 3 (N = 1049) was conducted on 5-6 August 2016.
The core question is whether these subtler primes also reduce affective polarization, which I measure here using the same-party and out-party party feeling thermometer ratings. If my theory is correct, I should find an increase in out-party feeling thermometer ratings, but not effect on same-party feeling thermometer ratings. Table 3 provides the results.

[Insert Table 3 about Here]

Table 3 confirms the findings in Table 1 from the main experiment: subjects who had their American identity primed feel more positively toward the other party, regardless of the type of prime used. Here, out-party feeling thermometer ratings increase by 3.6 degrees in experiment 2, and 3 degrees in experiment 3, and the effect is statistically significant in both cases.\textsuperscript{10} It is slightly smaller than the effects found in the main experiment above (about 35% smaller), but the key fact is that the main effect replicates in these different samples with these different primes. Further, as in the main study, there is no treatment effect on same-party feeling thermometers, which again is highly consistent with my underlying theoretical mechanism. Together, these results offer a nice robustness check on the main experimental findings above—multiple methods of priming American identity yield substantively similar results.

**Effects outside the Experimental Context**

The analysis above shows credibly that priming American identity makes subjects view the opposing party more positively due to a process of re-categorization. Such

\textsuperscript{10} Further, consistent with my theoretical account, in experiment 3, the stronger one identifies as an American, the larger the effect on out-party feeling thermometers.
experimental evidence is invaluable, but as with any experiment, there are concerns about how such effects generalize beyond the experimental context. Luckily, there is a real-world scenario that mimics the experimental stimulus by exogenously priming American national identity: the yearly July 4th holiday. The day is a celebration of America, and many Americans commemorate the day by displaying national symbols such as the American flag, attending parades, and singing patriotic songs. Additionally, newspapers often publish stories like the one used in experiment 1 above, which explain why people are proud to be Americans (see the online appendix for more on this point). As a result, Americans feel more American on that day (Madestam and Yanagizawa-Drott 2012).

Subjects who are interviewed around July 4th should have a slightly more positive impression of the other party and its leaders, all else equal. Finding evidence of this same effect in a non-experimental setting would be a strong buttress to my experimental results above.

If such effects exist, however, they are almost certainly quite small, as any shift in national identity would be modest and short-lived. To detect these effects, I therefore need an especially large dataset with considerable variation in date of interview. Luckily, the 2008 Annenberg National Election Study’s online panel provides me with just such a dataset: over the course of the 2008 campaign, nearly 30,000 subjects were interviewed between one and five times and asked a variety of political items. The data are a high-quality, nationally representative sample collected by Knowledge Networks (now GfK Custom Research). What makes the Annenberg data particularly appealing for my purposes, however, is that the interview date is randomly assigned within wave, so that there should be no systematic differences between the subjects who respond to the survey
near July 4th and those who answer it at other times (for more details on the study, see Johnston 2008). This data allows me the rare opportunity to test my theoretical claim in a non-experimental setting.

More specifically, I use the responses from wave 3 of the 2008 Annenberg panel, which interviewed subjects interviewed between April and August 2008. Because the Democratic Party primary in 2008 ran until early June, responses before June could be different from those that come afterwards (due to learning about the candidates in the Democratic primary). To avoid this potential confound, I focus my analyses on those interviewed between June and August. Unfortunately, the NAES does not include party feeling thermometers, but they did ask 100-point feeling thermometers about Obama and McCain, the two major-party nominees. So I examine how Democrats (Republicans) rate McCain (Obama) on the feeling thermometer scale, and whether those responses are more positive for those interviewed near July 4th. As above, I pool Democrats and Republicans together into one analysis, as I find no clear pattern of partisan differences in the data.

My expectation is that those interviewed “close” to July 4th will evaluate the other party’s nominee more positively. But what constitutes “close”? The answer is unclear. To address this issue, I use multiple definitions of “close.” First, I compute the number of days/weeks between the interview date and July 4th as a measure of closeness, with the expectation that as the number of days/weeks increases, the feeling thermometer rating of

---

11 Because I use the wave 2 data as well, my sample is the 14,524 respondents who were interviewed in both waves.

12 As I discuss in the online appendix, party and candidate feeling thermometer ratings are highly correlated, and this difference is very unlikely to affect my substantive results (see also Lelkes et al. 2017).
the opposing party’s nominee should decrease. Second, I look at those interviewed in various “windows” around July 4th, and compare them to those interviewed at similar points in June and August. For example, I compare those interviewed on July 4th (the first Friday in July) to those interviewed on either the first Friday in June or August. This is a one-day window around July 4th. I make the same comparison for a 3-day window, a 5-day window, a 7-day window, a 10-day window, and a 14-day window around July 4th. I expect the effect will dissipate as the window gets larger, as it includes more people interviewed at a greater distance from July 4th. Table 4 presents the results across this range of specifications.

[Insert table 4 about here]

While there is some variation in the size of the effect, the results in table 4 all point to a similar substantive conclusion: those interviewed close to July 4th have a more positive evaluation of the opposing party’s nominee. Those who are interviewed further away from July 4th have a lower feeling thermometer rating of the opposing party, and those interviewed close to July 4th have a higher rating. This is true of both the direct measure of closeness (the number of days/weeks since July 4th), and for the measures based on the windows around July 4th. For example, those interviewed in the 14-day window around July 4th rate the opposing party’s nominee 1.8 degrees warmer than those interviewed at otherwise similar periods in early June or August. Likewise, looking at the 5 day window around July 4th, feeling thermometer ratings there are approximately 2.7 degrees warmer than the similar periods in June/August, and in the 3-day window, ratings are 3.6 degrees warmer. In general, the effects get larger as the window gets more narrow, though in the limit—when we consider those who responded on July 4th itself—
the sample is too small to detect the effect (only N=86 people were interviewed that day), though it is properly signed. In short, proximity to July 4\textsuperscript{th} increased feeling thermometer ratings for the opposition.

But what if the effect is not due to the priming of American national identity that accompanies July 4\textsuperscript{th}, but rather to some other process occurring at the same time? There is no perfect way to guard against this, but there is one placebo test that I can run to help rule out this possibility. As I explained above, priming American identity should only affect opposite party candidate evaluations, not same party candidate evaluations. Same party candidates are part of both the national and partisan in-group, so increased in-group identity should not change how respondents feel about them. If I run the parallel set of regressions for same-party candidate feeling thermometers, I find null effects: proximity to July 4\textsuperscript{th} does not increase same-party feeling thermometer ratings (see the online appendix). This suggests that it really is American identity—and not some other factor—that generates the effects.

I can also guard against potential unobserved heterogeneity by using the feeling thermometer ratings from wave 2 (January-March 2008) as a control variable. This allows me to control for prior affect toward the candidates, to ensure that there was not some pre-existing difference between the groups that would explain the results. This is a very restrictive test: given that overall feeling thermometer scores correlate at nearly 0.7 between these two waves, there is little wave-to-wave change, and the changes found above were only on the order of 2-3 degrees. So finding any effect of proximity to July 4\textsuperscript{th} controlling for previous score is remarkable. In the rightmost column in Table 4, we see that the effects of days/weeks away from July 4\textsuperscript{th} survive even controlling for the
feeling thermometer score from the previous wave. The results for the various-sized windows become insignificant, which is not ideal, but again, the test is so restrictive that any finding here is surprising. This further bolsters the claim that this is a genuine effect.

Of course, this alone does not rule out some additional sort of unobserved heterogeneity: perhaps those interviewed close to July 4th are simply different in some other way. The research design itself guards against this possibility—because interview dates are randomly assigned within the wave, there should be nothing systematically different about those interviewed near July 4th than those interviewed at other times. However, I can also control for various demographic differences, and the results are unaffected (see the online appendix). I can also use matching methods to control for any pre-existing differences, and I still find the same substantive results: being interviewed near July 4th boosts out-party feeling thermometers (again, see the online appendix).

Overall, then, there is compelling evidence that the July 4th holiday temporarily reduces hostility toward the opposing party, consistent with my theoretical account. Not only is there survey experimental evidence of this effect, there is also real-world evidence as well.\(^{13}\)

**Conclusion and Implications**

This paper considers the ability of a subtle prime—increasing individuals’ sense of American national identity—to reduce affective polarization. Based on the logic of the

\(^{13}\) Additionally, as I show in the online appendix, I find similar effects during the 2008 summer Olympics, another event that boosts national identity. This is also consistent with Carlin and Love (Forthcoming), who report that partisan animus decreases in the period immediately following the death of Osama Bin Laden.
Common Ingroup Identity Model, I argue that when respondents see those from the other party as fellow Americans—rather than as members of a separate political tribe—affective polarization (that is, dislike and distrust of the opposing party) will be lower. Using both several original experiments, and a natural experiment stemming from the July 4th national holiday, I find strong support for my argument. I also find that such effects occur broadly, taking place in both parties, and across lines of partisan strength, as well as across race and gender. Many Americans—not just a small sub-group—respond to these types of national identity primes. Such results have the potential to change how many Americans see the opposition, at least in the short term.\textsuperscript{14}

More broadly, this study has important implications for our understanding of how scholars and practitioners might mitigate polarization in the United States. While affective polarization can spur on greater activism and political activity (Huddy et al. 2015), it has more pronounced negative effects by exacerbating gridlock and dissensus in our governing institutions (Hetherington and Rudolph 2015). But by a relatively subtle shift in rhetoric—characterizing the other side as fellow Americans rather than as a separate political tribe—dislike of the opposition can be softened. This work therefore establishes that it is possible to mitigate this sort of discord.

But this raises an important question: if these primes are successful, why don’t politicians use this strategy? The short answer is that they do use them, and they have a

\textsuperscript{14}These results, like those in Transue (2007), are largely positive, suggesting that strengthening American identity can make others behave in desirable ways (to support more education spending, or to be less affectively polarized). But as Theiss-Morse (2009) shows, stronger national identity can also carry negative consequences as well. An important topic for future work is understanding when such appeals have positive vs. negative effects.
long history in American politics dating back to Thomas Jefferson’s famous remark at his first inaugural address that “We are all Federalists, we are all Republicans.” Since then, many leaders have invoked the idea of transcending partisanship and coming together as Americans, from Truman’s call to put country ahead of party, to Reagan’s call to unite not as partisans but Americans in common cause, to President Obama’s repeated calls to bridge the partisan divide in Washington. For example, President Obama invoked this logic after the 2013 government shutdown, remarking that: “We come from different parties, but we are Americans first…Our regard for them [the American people] compels us all, Democrats and Republicans, to cooperate, and compromise, and act in the best interests of our nation — one nation, under God, indivisible with liberty and justice for all” (Obama 2013). Partisan elites do try to prime American identity in their rhetoric.

But the very nature of these appeals highlights why they often fail for politicians. When a politician makes such an appeal, especially in an age of elite polarization, voters will most likely see it through a partisan lens (Nicholson 2012). Indeed, while both Presidents Bush and Obama tried to transcend partisan differences and unite Americans, they both failed to do so. Ironically, because the president—or any politician—is seen as a partisan figure, she or he has a limited ability to successfully make such appeals. While non-partisan efforts like No Labels (Tuhoy 2015) could possibly work to overcome this limitation, they have struggled to gain traction in a highly polarized climate.

Ultimately, my results here suggest that a non-political strategy to prime American identity will be more effective. For example, in the real world, having more stories like those used in experiment 1—stories that stress what Americans have in common, and the positive aspects of American identity—would likely do more to
mitigate affective polarization. Establishing who can credibly issue such cues, under what conditions, and how long such effects last is an important next step in this sort of the research.  

No simple psychological prime can ever be a panacea for a complex phenomenon like affective polarization—if it could be, someone would have already done it. But that said, these results highlight a mechanism for reducing affective polarization, at least temporarily. But given the persistence and significance of this discord, these findings are an important first step of reducing this animosity.

15 Another important dimension to future research will be exploring partisan heterogeneity in treatment effects. While both Democrats and Republicans respond similarly here—because both share and value their American identity—in other contexts, that need not be true (e.g., Grossman and Hopkins 2016).
References


Huddy, Leonie. 2001. “From Social to Political Identity.” *Political Psychology* 22(1)


Biographical Sketch: Matthew S. Levendusky is associate professor of Political Science at the University of Pennsylvania, Philadelphia, PA, 19104.

Acknowledgements: Thanks to Marc Hetherington, Dan Hopkins, Yph Lelkes, Michele Margolis, Neil Malhotra, Yotam Margalit, Marc Meredith, Deborah Schildkraut, Omer Yair, the editors and anonymous referees, and participants at the 2016 ISPP and APSA Meetings for helpful comments and suggestions.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Out-Party FT</th>
<th>Same-Party FT</th>
<th>Obama FT (R)</th>
<th>Obama FT (D)</th>
<th>Trait Index</th>
<th>Positive Traits</th>
<th>Negative Traits</th>
<th>American Trait</th>
<th>Likes</th>
<th>Dislikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>5.599</td>
<td>1.646</td>
<td>4.216</td>
<td>2.651</td>
<td>0.155</td>
<td>0.224</td>
<td>-0.0038</td>
<td>0.209</td>
<td>0.0928</td>
<td>0.0953</td>
</tr>
<tr>
<td></td>
<td>(1.141)</td>
<td>(1.218)</td>
<td>(2.187)</td>
<td>(1.80)</td>
<td>(0.0378)</td>
<td>(0.0465)</td>
<td>(0.0525)</td>
<td>(0.0565)</td>
<td>(0.0458)</td>
<td>(0.0889)</td>
</tr>
<tr>
<td></td>
<td>[3.72,7.48]</td>
<td>[-0.36,3.65]</td>
<td>[0.61,7.82]</td>
<td>[-0.35,5.63]</td>
<td>[0.09,0.21]</td>
<td>[0.14,0.30]</td>
<td>[-0.09,0.08]</td>
<td>[0.12,0.30]</td>
<td>[0.02,0.17]</td>
<td>[-0.05,0.24]</td>
</tr>
<tr>
<td>Constant</td>
<td>25.77</td>
<td>65.24</td>
<td>22.63</td>
<td>65.15</td>
<td>2.633</td>
<td>2.613</td>
<td>3.493</td>
<td>3.325</td>
<td>0.391</td>
<td>1.562</td>
</tr>
<tr>
<td></td>
<td>(0.785)</td>
<td>(0.836)</td>
<td>(1.508)</td>
<td>(1.237)</td>
<td>(0.0259)</td>
<td>(0.0319)</td>
<td>(0.036)</td>
<td>(0.0388)</td>
<td>(0.0314)</td>
<td>(0.061)</td>
</tr>
<tr>
<td></td>
<td>[24.5,27.1]</td>
<td>[63.9,66.6]</td>
<td>[20.1,25.1]</td>
<td>[63.1,67.2]</td>
<td>[2.6,2.7]</td>
<td>[2.6,2.7]</td>
<td>[3.4,3.6]</td>
<td>[3.3,3.4]</td>
<td>[0.3,0.4]</td>
<td>[1.5,1.7]</td>
</tr>
</tbody>
</table>

| N         | 1,695        | 1,711         | 694          | 1,005        | 1,716       | 1,716          | 1,716          | 1,716         | 1,729  | 1,729    |
| R-squared | 0.014        | 0.001         | 0.005        | 0.002        | 0.01        | 0.013          | 0              | 0.008         | 0.002  | 0.001    |

Table 1: Effects of American Prime on Out-Party Animus, Experiment 1

*Note:* cell entries are OLS regression coefficients with associated standard errors in parentheses and 90% confidence intervals in brackets. Treatment effects that can be differentiated from 0 at conventional levels ($\alpha < 0.10$, two-tailed) are given in **bold**.
<table>
<thead>
<tr>
<th></th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>4.296</td>
<td>6.318</td>
<td>2.915</td>
<td>3.327</td>
<td>0.153</td>
<td>0.196</td>
<td>0.0953</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>(1.438)</td>
<td>(1.667)</td>
<td>(2.686)</td>
<td>(3.097)</td>
<td>(0.048)</td>
<td>(0.059)</td>
<td>(0.058)</td>
<td>(0.07)</td>
</tr>
<tr>
<td></td>
<td>[1.93, 6.66]</td>
<td>[3.57, 9.06]</td>
<td>[-1.51, 7.34]</td>
<td>[-1.78, 8.43]</td>
<td>[0.07, 0.23]</td>
<td>[0.10, 0.29]</td>
<td>[-0.002, 0.2]</td>
<td>[-0.02, 0.22]</td>
</tr>
<tr>
<td>Strong Partisan</td>
<td>-8.891</td>
<td>-5.753</td>
<td>-0.134</td>
<td>-0.181</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.595)</td>
<td>(3.148)</td>
<td>(0.0531)</td>
<td>(0.0645)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-11.5, -6.3]</td>
<td>[-10.9, -5.57]</td>
<td>[-0.22, -0.05]</td>
<td>[-0.29, -0.07]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment* Sorted</td>
<td>3.425</td>
<td>3.479</td>
<td>0.0065</td>
<td>-0.0054</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.319)</td>
<td>(4.618)</td>
<td>(0.0774)</td>
<td>(0.0939)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-0.40, 7.24]</td>
<td>[-4.12, 11.09]</td>
<td>[-0.12, 0.13]</td>
<td>[-0.16, 0.15]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorted</td>
<td>-9.932</td>
<td>-20.07</td>
<td>-0.199</td>
<td>-0.0569</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.598)</td>
<td>(2.983)</td>
<td>(0.0536)</td>
<td>(0.0671)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-12.6, -7.3]</td>
<td>[-25, -15.2]</td>
<td>[-0.29, -0.11]</td>
<td>[-0.17, 0.05]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment* Sorted</td>
<td>-2.343</td>
<td>-0.366</td>
<td>-0.0946</td>
<td>-0.0112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.327)</td>
<td>(4.326)</td>
<td>(0.0782)</td>
<td>(0.0978)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-6.17, 1.49]</td>
<td>[-7.49, 6.76]</td>
<td>[-0.22, 0.03]</td>
<td>[-0.17, 0.15]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>29.18</td>
<td>30.95</td>
<td>24.67</td>
<td>33.71</td>
<td>2.685</td>
<td>2.743</td>
<td>0.459</td>
<td>0.428</td>
</tr>
<tr>
<td></td>
<td>(0.988)</td>
<td>(1.159)</td>
<td>(1.874)</td>
<td>(2.179)</td>
<td>(0.0329)</td>
<td>(0.0388)</td>
<td>(0.0398)</td>
<td>(0.0485)</td>
</tr>
<tr>
<td></td>
<td>[27.55, 30.81]</td>
<td>[29.04, 32.85]</td>
<td>[21.6, 27.75]</td>
<td>[30.1, 37.3]</td>
<td>[2.63, 2.74]</td>
<td>[2.67, 2.81]</td>
<td>[0.39, 0.52]</td>
<td>[0.35, 0.51]</td>
</tr>
</tbody>
</table>

Table 2: Testing for Heterogeneous Treatment Effects, Sorted & Strong Partisans, Experiment 1

*Note:* cell entries are OLS regression coefficients with associated standard errors in parentheses and 90% confidence intervals in brackets. Treatment effects that can be differentiated from 0 at conventional levels are given in **bold.**
<table>
<thead>
<tr>
<th></th>
<th>(1) Out-Party FT Experiment 2</th>
<th>(2) Same-Party FT Experiment 2</th>
<th>(3) Out-Party FT Experiment 3</th>
<th>(4) Same-Party FT Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3.56</td>
<td>0.19</td>
<td>3.06</td>
<td>-0.48</td>
</tr>
<tr>
<td></td>
<td>(1.96)</td>
<td>(1.99)</td>
<td>(1.20)</td>
<td>(1.69)</td>
</tr>
<tr>
<td></td>
<td>[0.33,6.81]</td>
<td>[-3.09,3.47]</td>
<td>[1.08,5.04]</td>
<td>[-3.26,2.3]</td>
</tr>
<tr>
<td>Constant</td>
<td>29.79</td>
<td>55.36</td>
<td>18.67</td>
<td>62.72</td>
</tr>
<tr>
<td></td>
<td>(1.70)</td>
<td>(1.72)</td>
<td>(0.86)</td>
<td>(1.20)</td>
</tr>
<tr>
<td></td>
<td>[26.99,32.58]</td>
<td>[52.53,58.19]</td>
<td>[17.25,20.08]</td>
<td>[60.73,64.70]</td>
</tr>
<tr>
<td>N</td>
<td>724</td>
<td>726</td>
<td>1,024</td>
<td>1,024</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.005</td>
<td>0.04</td>
<td>0.006</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 3: Effects of American Prime on Out-Party Feeling Thermometers, Experiments 2 & 3

*Note: cell entries are OLS regression coefficients with associated standard errors in parentheses and 90% confidence intervals in brackets. Treatment effects that can be differentiated from 0 at conventional levels are given in bold.*
Table 4: The Effects of the Natural Experiment of July 4th on Feeling Thermometer Ratings of the Opposing Party Nominee

Note: Cell entries show the effect of being interviewed close to July 4th on the feeling thermometer ratings for the opposing party’s nominee using the 2008 Annenberg National Election Study online panel. Estimates are coefficient estimates, with associated standard errors in parentheses and 90% confidence intervals in brackets. Estimates that can be differentiated from 0 at conventional levels are given in bold.