

WHEN EFFORTS TO DEPOLARIZE THE ELECTORATE FAIL

MATTHEW S. LEVENDUSKY*

Abstract The mass public has become affectively polarized—ordinary Americans increasingly dislike and distrust those from the other party, with negative consequences for politics. Drawing on work in political and social psychology, this paper tests two mechanisms for reducing this discord, both of which have been shown to reduce similar biases in other settings: heightening partisan ambivalence, and using self-affirmation techniques. A population-based survey experiment shows that neither strategy reduces affective polarization in the aggregate. But this null finding masks an important heterogeneity: Heightening partisan ambivalence reduces affective polarization for ideological moderates, but increases such discord for those with more extreme ideological identities. Efforts to depolarize the electorate can make it more deeply divided, with important implications for our understanding of contemporary politics and the durability of affective polarization.

Few public opinion topics have attracted more attention in recent years than political polarization. While scholars disagree on whether ordinary voters' issue positions have become more divided (Fiorina, Abrams, and Pope 2005), a consensus exists that affective polarization has increased—voters today dislike and distrust those from the other party more than they did just a few decades ago (Iyengar, Sood, and Lelkes 2012; Abramowitz and Webster 2016). Such affective polarization lowers political trust and makes governance more difficult (Hetherington and Rudolph 2015), hampers interpersonal relations

MATTHEW S. LEVENDUSKY is a professor in the Department of Political Science (and, by courtesy, in the Annenberg School for Communication) at the University of Pennsylvania, Philadelphia, PA, USA. The author thanks the following individuals and groups for helpful comments on this paper: the editors and anonymous referees, as well as Jamie Druckman, Don Green, Bryan Gervais, the Penn American Politics Working Group, seminar participants at Columbia University, and seminar participants at the 2017 Annual Meeting of the Midwest Political Science Association. Data were collected by Time-Sharing Experiments for the Social Sciences [NSF Grant Number 1227179 to Jeremy Freese and James Druckman, Principal Investigators]. *Address correspondence to Matthew S. Levendusky, Department of Political Science, 133 South 36th Street, Philadelphia, PA 19104, USA; email: mleven@sas.upenn.edu.

doi:10.1093/poq/nfy036

© The Author(s) 2018. Published by Oxford University Press on behalf of the American Association for Public Opinion Research. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

(Huber and Malhotra 2017), and hinders economic exchange (McConnell et al. 2018).

Is it possible to reduce this affective polarization? As it stems primarily from the power of partisanship to divide the world into ingroups and outgroups, weakening this group-centric thinking should reduce affective polarization. This paper outlines and tests two mechanisms for achieving this goal: increasing partisan ambivalence (Lavine, Johnston, and Steenbergen 2012) and using self-affirmation techniques (Steele 1988).

Using a population-based survey experiment, I find no evidence that the self-affirmation treatment has any effect on affective polarization. Heightening partisan ambivalence similarly has no aggregate effect, but this masks an important heterogeneity: The treatment decreases affective polarization for ideological moderates, but increases such polarization for others. Moderates enter the study much less polarized than other respondents, so the treatment works best for those who need it least, and does little to actually reduce discord in the electorate. These results highlight that reducing affective discord will likely be quite difficult, precisely because of its intimate connection to our most fundamental political identity: partisanship.

Affective Polarization, Social Identity, and Efforts to Mitigate It

Affective polarization refers to a dislike of members of the opposing party, and a tendency to impute negative characteristics to them (Iyengar, Sood, and Lelkes 2012, p. 407).¹ For example, the average feeling-thermometer ratings given to the opposing party have dropped considerably since the late 1980s (Iyengar, Sood, and Lelkes 2012), with 2016 representing the apex of such animus toward the other party (Pew Research Center 2016). For some individuals, partisan prejudice now outstrips racial prejudice (Iyengar and Westwood 2015). Such affective polarization is not simply cheap talk—individuals will pay a real economic cost to signal their partisan allegiance (McConnell et al. 2018), and this has effects not only on governance (Hetherington and Rudolph 2015), but also on everyday interactions (Huber and Malhotra 2017).

Affective polarization stems directly from identification with a political party. Identifying with a party divides the world into an ingroup (one's own party) and an outgroup (the opposing party), which leads to ingroup favoritism, the tendency to like one's own (partisan) group and dislike the outgroup (Tajfel and Turner 1979). Affective polarization is an outgrowth of this classification: Because the other party is an outgroup, they take on a negative connotation and individuals dislike and distrust them (Iyengar, Sood, and

1. Note that affective and ideological (issue-based) polarization are distinct concepts that are only weakly related empirically (Mason 2015)—one can increase while the other decreases (Levendusky and Malhotra 2016).

Lelkes 2012; Iyengar and Westwood 2015). The contemporary political environment—with polarized elites starkly divided along partisan lines—further accentuates this tendency toward partisan animus and affective polarization.

This definition of affective polarization also suggests a mechanism for mitigating it: weakening partisan group-centrism and accompanying biases. Here, I draw on two strategies that have ameliorated similar biases in other contexts and test their ability to lessen affective polarization. If strategies that have worked elsewhere fail here, it tells us something important about how ingrained affective polarization is, and how difficult it likely will be to overcome.

The first strategy involves heightening partisan ambivalence, a conflict between individuals' partisan identities and their short-term evaluations of or feelings toward the parties (Lavine, Johnston, and Steenbergen 2012, pp. 3–5). When individuals are univalent (nonambivalent) partisans, partisanship strongly colors their attitudes, vote choice, economic assessments, and evaluation of new information, precisely because partisanship is so central to political reasoning. But when ambivalence is induced (by, say, reflecting on what one dislikes about one's own party and likes about the other party), partisanship is displaced from its normally central role in political decision-making, and individuals lessen their reliance on it as a heuristic (Lavine, Johnston, and Steenbergen 2012; Bolsen, Druckman, and Cook 2014). Ambivalence, by suggesting where individuals diverge from their party, also weakens same-party identification, and the ability of partisanship to divide up the world into ingroups and outgroups (Brewer 1991). So, by ameliorating partisan group-centrism, ambivalence dampens affective polarization and discord. *Hypothesis 1: Priming partisan ambivalence will reduce affective polarization.*

Second, self-affirmation techniques can be used to the same ends. Self-affirmation techniques build from the recognition that affective polarization and related phenomena reflect a desire to maintain a positive self-image (Oakes and Turner 1980). Individuals want to think highly of themselves, and favoring their partisan ingroup (and disparaging the outgroup) is a mechanism for achieving that goal. But of course there are also many apolitical mechanisms for maintaining a good self-image. If individuals reflect on, for instance, their musical talent, business acumen, or any other positive apolitical trait, they feel good about themselves and have less need to defend their partisan group, thereby reducing partisan bias and affective polarization (Steele 1988). For example, Binning et al. (2010) show that, around the 2008 election, self-affirmation reduced partisan biases in candidate evaluations. Similarly, numerous studies show that self-affirmation makes individuals more open to new information, and less defensive of their side's positions (e.g., Cohen et al. 2007). This same logic applies to affective polarization: Affirming some apolitical aspect of an individual's personality lessens their need to engage in identity-protecting group favoritism, so they will have less need to disparage the other party. *Hypothesis 2: Self-affirmation techniques will reduce affective polarization.*

However, there should be an important heterogeneity in the effectiveness of these treatments. Respondents with stronger partisan or ideological identities have more efficacious and durable identities that will resist the impact of these types of primes (Mason 2015), and hence the treatment effects will be smaller for these individuals. *Hypothesis 3: Both self-affirmation and partisan ambivalence will be less effective at reducing affective polarization for those with stronger identities.*

Experimental Strategy and Results

I investigate whether either strategy—increasing partisan ambivalence or using self-affirmation techniques—can reduce affective polarization using a population-based survey experiment. Subjects first provide their partisanship, then are randomized into one of three conditions: an apolitical control condition, the partisan ambivalence condition, or the self-affirmation condition. To induce partisan ambivalence, I adopt the strategy from Lavine, Johnston, and Steenbergen (2012) and ask subjects to list two things they dislike about their own party and two things they like about the other party (pp. 92–93). For the self-affirmation treatment, I follow McQueen and Klein (2006) and ask subjects to pick a characteristic or value that is important to them from a long list of options (i.e., creativity, athletic talent, etc.). They then write a brief paragraph explaining a time when they demonstrated that characteristic/value. Full treatment stimuli are provided in the online [appendix](#).

Post-treatment, subjects answer items designed to measure affective polarization. Because there is not one single measure used by earlier studies, the study included three different items from the literature: a feeling-thermometer rating of the other party (Hetherington and Rudolph 2015; Abramowitz and Webster 2016); social distance measures (here, whether the respondent is comfortable having close friends from the other party; see Iyengar, Sood, and Lelkes [2012]); and whether the other party's ideas are so extreme they are dangerous for the health of the nation (Pew Research Center 2016). The last item is a more severe manifestation of affective polarization, but in the Pew data, this was not an uncommon sentiment—approximately 40 percent of Democrats and Republicans agreed with that statement in the 2016 Pew survey.

Respondents also state their preferred degree of partisan homogeneity when discussing politics: Would they prefer a group that is almost all members of their own party, or one that is more mixed, with members of both parties (Klar 2014)? While this is not a measure of affective polarization per se, such discussion preferences are a consequence of it, as high levels of affective polarization lead people to avoid interacting with the other party (Lelkes and Westwood 2017). Reducing affective polarization therefore should increase willingness to engage in cross-party discussion, which has positive spillover benefits, as such conversations promote tolerance for opposing viewpoints (Mutz 2006). The full text of all items appears in the online [appendix](#).

The experiment was conducted by GfK Custom Research (previously Knowledge Networks), which uses random digit dialing and address-based sampling to recruit representative samples of the US population for studies such as this one. A total of 2,095 respondents completed the survey on April 15–25, 2016, for a completion rate of 55 percent and an AAPOR Response Rate 3 of 4.8 percent. The study was conducted through Time Sharing Experiments for the Social Sciences (TESS); the TESS proposal—which serves as the pre-analysis plan for this experiment—is provided in the online [appendix](#).

Table 1 shows the effect of both treatments on the measures of affective polarization as well as the discussion preference item. The results are disappointing: Neither treatment has any significant effect on any of the outcome measures of affective polarization, and neither has any effect on preferences for homogeneous versus heterogeneous discussion.²

The effects above are the intent-to-treat effects: They show the effect of being assigned to the treatment condition, regardless of whether the respondent actually complied with the treatment. A closer look at the data reveals that many subjects failed to comply with the partisan-ambivalence treatment. When asked about what they like about the opposing party, many subjects

Table 1. Effects of treatments on affective polarization

	(1) Other party feeling thermometer	(2) Other party is a danger	(3) Friends with other party	(4) Mixed political discussion
Ambivalence treatment	1.47 (1.18)	0.05 (0.05)	-0.03 (0.04)	0.06 (0.06)
Self-affirmation treatment	0.24 (1.18)	0.01 (0.05)	0.04 (0.04)	0.04 (0.06)
Constant	24.91** (0.82)	3.93** (0.04)	3.22** (0.03)	3.16** (0.04)
Observations	2,025	2,071	2,035	2,011
R-squared	0.00	0.00	0.00	0.00

NOTE.—Cell entries are OLS regression coefficients with associated standard errors in parentheses; see the text for details on the dependent variable measures.

** $p < 0.01$ (one-tailed)

2. One concern with null results is that the manipulation failed (i.e., the treatment did not increase ambivalence or affirm apolitical dimensions of individuals' identities). I discuss the ambivalence treatment's effectiveness below. The online appendix includes a manipulation check item that shows the self-affirmation manipulation worked, but simply did not affect the dependent variables.

wrote “nothing” (or some similar variant, often with considerable emphasis), or they responded with an answer like “they have a low chance of winning [the 2016 election],” “are you kidding me,” or “I like it being 50 years since I voted for a Democrat.” A research assistant coded the responses to determine whether respondents actually listed two valid likes/dislikes and found that only 38 percent did so. In contrast, 88 percent complied with the self-affirmation treatment, so it is not simply that subjects will not respond to any prompt. Rather, the hostility toward the ambivalence treatment is an important indicator of the degree of affective polarization of the mass public in and of itself.³

Compliance with the ambivalence treatment is likely to be nonrandom, and those with stronger identities should be less likely to comply with it—their more crystallized identities make them less willing to consider the positive dimensions of the other party. This is the case: Strong partisans are less likely—and ideological moderates more likely—to comply with this treatment. This finding suggests that there should be heterogeneous treatment effects generated by partisan and ideological identity strength.

As shown in [table 2](#), partisan strength does not moderate the effectiveness of either treatment, but ideological strength does moderate the effect of the ambivalence treatment.⁴ Take, for example, the feeling-thermometer item: There are null effects for those with stronger ideological identities, but for those with weaker ideologies (moderate or slightly liberal/conservative respondents), the treatment increases ratings by three degrees. Likewise, for these same moderate respondents, priming ambivalence makes them 30 percent less likely to rate the opposing party at the coldest rating (0 degrees). A similar pattern for group discussion preferences emerges, with an effect for ideological moderates but no effects for stronger ideologues: Ambivalence only makes moderate respondents prefer a more heterogeneous group.

There is one outcome, however, where the ambivalence treatment actually *increases* affective polarization for stronger ideologues. The ambivalence treatment increases the percentage of respondents who “strongly agree” that the other party’s policies are a threat to the nation; this figure jumps by 11 percentage points (or a 26 percent relative increase) in the ambivalence condition. In some cases, then, the treatment backfires and further heightens affective polarization.⁵

3. One limitation with the ambivalence treatment is that treatment failure (subjects refuse to consider something they like about the other party) cannot be disentangled from an ineffective treatment (subjects know what they like about the other party, but it has no effect). I thank an anonymous referee for making this point.

4. The treatment is less effective for voters with stronger identities partly because their dislike of the other party may be more rooted in policy, rather than idiosyncratic factors. I thank an anonymous referee for this point.

5. Heterogeneous treatment effects did not exist across party (Democrats vs. Republicans), age, gender, race, education, or income. Also, effects did not differ based on proximity to a presidential primary context. Results are presented in [tables A1–A7](#) of the online appendix.

Table 2. Heterogeneous treatment effects by partisan and ideological strength

	(1) Other party feeling thermometer	(2) Other party is a danger	(3) Friends with other party	(4) Mixed political discussion	(5) Other party feeling thermometer	(6) Other party is a danger	(7) Friends with other party	(8) Mixed political discussion
Ambivalence treatment	3.17* (1.51)	-0.04 (0.07)	0.001 (0.06)	0.11# (0.08)	2.20# (1.41)	0.04 (0.06)	-0.02 (0.05)	0.09 (0.07)
Self-affirmation treatment	0.52 (1.49)	-0.01 (0.07)	0.07 (0.06)	0.04 (0.08)	-0.38 (1.40)	-0.02 (0.06)	0.04 (0.05)	-0.03 (0.07)
Ideologue	-8.59** (1.62)	0.33** (0.07)	-0.12* (0.06)	-0.48** (0.08)				
Ambivalence*Ideologue	-3.91* (2.32)	0.21* (0.10)	-0.07 (0.09)	-0.10 (0.12)				
Self-affirmation*Ideologue	-1.08 (2.32)	0.07 (0.10)	-0.09 (0.09)	-0.02 (0.12)				
Strong partisan					-14.44** (1.63)	0.57** (0.07)	-0.17** (0.06)	-0.74** (0.08)
Ambivalence*Strong partisan					-0.84 (2.33)	-0.00 (0.10)	-0.01 (0.09)	-0.02 (0.12)
Self-affirmation*Strong partisan					2.31 (2.32)	0.07 (0.10)	-0.01 (0.09)	0.21** (0.12)
Constant	28.52** (1.05)	3.79** (0.05)	3.27** (0.04)	3.36** (0.05)	30.02** (0.97)	3.73** (0.04)	3.28** (0.04)	3.42** (0.05)
Observations	2,025	2,071	2,035	2,011	2,025	2,071	2,035	2,011
R-squared	0.06	0.05	0.01	0.05	0.10	0.09	0.01	0.09

NOTE.—Cell entries are OLS regression coefficients with associated standard errors in parentheses.
$p < 0.10$; * $p < 0.05$; ** $p < 0.01$ (one-tailed)

Such findings underline an important limitation of these treatments to change outcomes in the real world. To the extent that the treatments reduced affective polarization, they did so for those with weaker ideological identifications. But these individuals came into the study less affectively polarized than those with stronger ideological identifications. In the control condition, stronger ideologues rate the opposing party 10 degrees colder on the feeling-thermometer scale, and they are twice as likely to strongly agree that the other party's policies are a threat to the nation (relative to weaker ideologues). The ambivalence treatment works best for those who need it least, and hence such treatments are unlikely to notably change real-world levels of affective polarization.

General Discussion and Conclusions

Can we reduce affective polarization and discord in the mass public? This research note tests two strategies for doing so: increasing partisan ambivalence and self-affirmation techniques. Both of these strategies have been shown to reduce partisan biases similar to affective polarization in other contexts, suggesting that they should similarly work here. Neither strategy reduces affective polarization in the aggregate. The partisan ambivalence treatment decreases affective polarization for those who are more ideologically moderate, but these individuals enter the study less polarized in the first place. Further, on at least one measure, the ambivalence treatment actually exacerbates affective polarization for those with stronger identities. This suggests that these treatments will have, at best, a very modest effect on reducing real-world levels of affective polarization and discord.

Such findings have several important implications. First, on the methodological front, they show the importance of avoiding the file-drawer problem by publishing null results, especially for treatments that have been shown to work in other contexts (on the file-drawer problem in political science, see [Franco, Malhotra, and Simonovits \[2015\]](#)). For example, self-affirmation techniques have worked in other political contexts ([Binning et al. 2010](#)), but they failed here, likely due to the chronic accessibility of partisanship in the current polarized political climate.

More broadly, these findings highlight that affective polarization is durable and not easy to overcome in most settings. Perhaps this should not be terribly surprising, given its close linkage to partisanship, arguably the most powerful political identity.

To overcome these effects, scholars and practitioners will need to find more powerful primes. For example, priming a sense of American (rather than partisan) identity can reduce affective polarization ([Levendusky 2018](#)), as can correcting misperceptions about the parties ([Ahler and Sood 2018](#)). Finding other such techniques powerful enough to reduce this discord is an important topic for future scholarship.

Supplementary Data

Supplementary data are freely available at *Public Opinion Quarterly* online.

References

- Abramowitz, Alan, and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of U.S. Elections in the 21st Century." *Electoral Studies* 41:12–22.
- Ahler, Douglas, and Gaurav Sood. 2018. "The Parties in Our Heads." *Journal of Politics* 80: 964–81.
- Binning, Kevin, David Sherman, Geoffrey Cohen, and Kirsten Heitland. 2010. "Seeing the Other Side." *Analyses of Social Issues and Public Policies* 10:276–92.
- Bolsen, Toby, James Druckman, and Fay Lomax Cook. 2014. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36:235–62.
- Brewer, Marilynn. 1991. "The Social Self: On Being the Same and Different at the Same Time." *Personality and Social Psychology Bulletin* 17:475–82.
- Cohen, Geoffrey, David Sherman, Anthony Bastardi, Lillian Hsu, Michelle McGoey, and Lee Ross. 2007. "Bridging the Partisan Divide." *Journal of Personality and Social Psychology* 93:415–30.
- Fiorina, Morris, Samuel Abrams, and Jeremy Pope. 2005. *Culture War? The Myth of a Polarized America*. New York: Pearson Longman.
- Franco, Annie, Neil Malhotra, and Gabor Simonovits. 2015. "Underreporting in Political Science Survey Experiments." *Political Analysis* 23:306–12.
- Hetherington, Marc, and Thomas Rudolph. 2015. *Why Washington Won't Work*. Chicago: University of Chicago Press.
- Huber, Gregory A., and Neil Malhotra. 2017. "Political Homophily in Social Relationships." *Journal of Politics* 79:269–83.
- Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76:405–31.
- Iyengar, Shanto, and Sean Westwood. 2015. "Fear and Loathing Across Party Lines." *American Journal of Political Science* 59:690–707.
- Klar, Samara. 2014. "Partisanship in a Social Setting." *American Journal of Political Science* 58:687–704.
- Lavine, Howard, Christopher Johnston, and Marco Steenbergen. 2012. *The Ambivalent Partisan*. New York: Cambridge University Press.
- Lelkes, Yphtach, and Sean Westwood. 2017. "The Limits of Partisan Prejudice." *Journal of Politics* 79:485–501.
- Levendusky, Matthew. 2018. "Americans, Not Partisans." *Journal of Politics* 80:59–70.
- Levendusky, Matthew, and Neil Malhotra. 2016. "Does Media Coverage of Partisan Polarization Affect Political Attitudes?" *Political Communication* 33:283–301.
- Mason, Liliana. 2015. "I Disrespectfully Agree." *American Journal of Political Science* 59:128–45.
- McConnell, Christopher, Neil Malhotra, Yotam Margalit, and Matthew Levendusky. 2018. "The Economic Consequences of Partisanship in a Polarized Era." *American Journal of Political Science* 62:5–18.
- McQueen, Amy, and William Klein. 2006. "Experimental Manipulations of Self Affirmation: A Systematic Review." *Self and Identity* 5:289–354.
- Mutz, Diana. 2006. *Hearing the Other Side*. New York: Cambridge University Press.
- Oakes, Penelope J., and John Turner. 1980. "Social Categorization and Intergroup Behavior." *European Journal of Social Psychology* 10:295–301.
- Pew Research Center. 2016. "Partisanship and Political Animosity in 2016." Available at <http://pewrsr.ch/2kjLB4S>, accessed June 2017.

- Steele, Claude. 1988. "The Psychology of Self-Affirmation: Sustaining the Integrity of the Self." In *Advances in Experimental Social Psychology*, edited by Leonard Berkowitz, pp. 261–302. New York: Academic Press.
- Tajfel, Henri, and John Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel, pp. 33–47. Monterey, CA: Brooks/Cole Publishing Company.