The Consequences of Social Interaction on Outparty Affect and Stereotypes

Erin Rossiter

Department of Political Science
Washington University in St. Louis

February 21, 2020

Abstract

Americans increasingly dislike members of the opposite political party and associate negative stereotypes with them such as close-minded, mean, and hypocritical. Yet, Americans interact, whether talking about politics or not, with opposing party members in their everyday lives. How do social interactions across party lines impact the negative feelings and perceptions Americans hold for opposing party members? How might the consequences of social interactions that touch on politics differ from those that do not? Using an experiment that manipulates whether or not a pair of opposing party members engage in social interaction, and if so, whether they discuss a political or non-political topic, I assess the effect of social interaction on how partisans feel about and how they think about the outparty. I find that social interaction mitigates negative outparty affect and deters future use of negative stereotypes to describe the outparty. This positive effect holds for both non-political and political conversations. Surprisingly, I do not find evidence that the effect of political conversation is any less than the effect of non-political conversation on improving a partisan’s negative view of the outparty. These results provide new evidence that interparty social interaction, regardless of whether the conversation is politically-charged or not, can work to undo the negative view of outparty members held by many Americans.

*Valuable feedback for this project was provided by Political Data Science Lab members at Washington University in St. Louis. I thank participants at APSA 2019 for helpful comments. Funding for this project was provided by the National Science Foundation (SES-1558907) and the Weidenbaum Center on the Economy, Government, and Public Policy at Washington University in St. Louis. I appreciate any comments or suggestions. Please do not cite or distribute without permission.
1 Introduction

Americans increasingly dislike members of the opposite political party and associate negative stereotypes with them such as close-minded, mean, and hypocritical. Yet, Americans interact, whether talking about politics or not, with opposing party members in their everyday lives. How do these interactions across party lines impact the negative feelings and perceptions Americans hold for opposing party members? When might social interaction with an opposing party member reinforce, and when might it mitigate, the negative view of outparty members held by many Americans? How might the consequences of social interactions that delve into political topics differ from those that do not?

Scholars often understand negativity toward the outparty through the framework of partisanship as a social identity. As a social identity, self-perceived membership in a political party leads Americans to categorize the world into the "ingroup" (their own party) or the "outgroup" (the opposing party) (Tajfel and Turner 1979), and this categorization impacts the way Americans feel and the way they think about the outparty. Negative feelings and perceptions about the outparty stem from a motivation to feel positively about ourselves and our ingroup, which is accomplished through direct comparison to the outgroup, or the outparty. As the country continues to divide along party lines, Americans’ partisan identities grow stronger and outparty negativity intensifies (Mason 2018).

There are two plausible expectations concerning how a negative view of the outparty might change after interacting with an outparty member. One expectation, stemming from the increasing strength of Americans’ partisan identities, is that exposure to outparty members will cue negative outparty feelings and perceptions, leading to an interaction that does not improve, or even worsens, an already negative view of the outparty. A second expectation stems from the literature on intergroup contact (e.g., Allport 1954), which finds that social interaction tends to be an effective strategy for improving negative outgroup attitudes for other social groups, so the idea is that this strategy should apply to partisan groups, as well.

While it may seem as though social identity theory and theories of intergroup contact are in tension when it comes to explaining the effects of partisan interaction, I argue that these theories are actually more complementary than is currently appreciated in the political science literature (e.g.,
Bond, Shulman and Gilbert 2018). Because social identity theory explains why we view outparty members negatively, it is a useful basis for understanding how these feelings and perceptions might change via contact. And by considering these theories in tandem, we can consider how the consequences of social interaction may differ if partisans talk about family, friends, and hobbies, or if partisans talk about politics.

Specifically, our social identities help us initially make sense of, and decrease uncertainty surrounding, any social interaction. When interacting with an outparty member, we initially categorize them as such, and negative perceptions and stereotypes about the outparty fill in for what we do not yet know. Contact, however, can alter a partisan’s representation of "outparty member." Because interactions with outparty members that occur in our everyday lives are usually informal, impromptu, and lack any particular goal, these interactions likely feature the sharing of personal information that allows opposing partisans to view each other as individual people rather than outparty prototypes. Therefore, social interactions personalize the outgroup member, breaking down the understanding of the outgroup member by the negative terms which define the outparty (Brewer and Miller 1984). In doing so, I expect that outparty contact undermines the usefulness of relying on partisan identity to make sense of future social interactions and interferes with the use of a biased, oversimplified representation of the outparty, generalizing the benefits of contact to the outparty at large.

A positive social interaction is fairly easy to imagine when conversation avoids overtly political topics—talking about work, family, or hobbies. These conversations are likely to promote an understanding of opposing party members on a personal level, replacing partisanship as the most useful lens to view them through. However, it is harder to imagine conversation that delves into political topics having a similar effect. It is unclear if political conversation provides an environment for sharing information that allows opposing party members to view each other as anything other than "the outparty." Nonetheless, I expect that positive, personalized interactions can occur even when conversation drifts into political topics, such that political conversation across party lines can improve negative outparty affect and stereotypes as well.

In this paper, I test these claims using an experiment in which I manipulate whether or not a pair of opposing party members engage in social interaction, and if so, whether they discuss a non-political or political topic. I am interested in two main outcomes: how partisans feel and
how they think about the outparty after social interaction with an outparty member. I find that interparty interaction mitigates negative outparty affect and deters future use of negative stereotypes to describe the outparty. This positive effect holds for both non-political and political conversations. Surprisingly, I do not find evidence that the effect of political conversation is any less than the effect of non-political conversation on improving negative outparty affect and perceptions. These results provide new evidence that interparty social interaction, regardless of whether the conversation is politically-charged or not, can work to undo the negative representation of outparty members held by many Americans. I conclude with a discussion of when we might expect partisans to engage in interparty interactions in their everyday lives, scope conditions that describe when interparty social interactions are likely to have these effects, and opportunities for future research.

2 How Americans feel and think about outparty members

The American political climate is characterized by a pronounced non-ideological division across party lines that is often said to take the general form of a "dislike and distrust" for members of the opposing party (see Iyengar et al. 2019). Research shows this trend manifests in many ways. There is a well-documented affective response toward outparty members—Republicans and Democrats increasingly report feeling negative toward members of the outparty (e.g., Iyengar, Sood and Lelkes 2012). Additionally, there is a cognitive response toward outparty members—partisans hold a negative, over-generalized representation of the outparty. This outparty representation includes negative trait stereotypes (Iyengar, Sood and Lelkes 2012), an overestimation of the extent to which outparty members belong to groups stereotypically associated with the outparty (Ahler and Sood 2018), and an overestimation of the extremity of opposing party members’ political views (Levendusky and Malhotra 2015).

One explanation for this general trend is a partisan-ideological sorting—conservatives increasingly identify as Republican and liberals increasingly identify as Democrat (Levendusky 2009; Mason 2015). Not only have ideological and partisan identities aligned, but race, religion, and more have sorted along the same partisan divide (Mason 2018). As identities that cut across party lines have decreased, the strength of Americans’ partisan identities has increased, which has affective and cognitive consequences. Stronger partisans react with stronger emotion to perceived party threats,
regardless of their ideological positions (Mason 2015, 2018). And, as clearer social distinctions are made between the parties and as Americans hold stronger partisan identities, it becomes easier and more tempting to make (potentially inaccurate) generalizations about the outparty (Westfall et al. 2015).

Talking across party lines has repeatedly been cited as a solution to America’s deep, bitter partisan divide. Not only do media (e.g., Grumet 2019) and politicians (e.g., Fang 2017) offer this advice, but nonprofits spend a great deal of money promoting this philosophy. But because the majority of political science research on the consequences of interparty social interaction focuses on, broadly, the sharing of political information (e.g., Berelson et al. 1954) and political participation (e.g., Huckfeldt and Sprague 1991; McClurg 2003; Mutz 2006), the consequences of talking with the political opposition on how Americans feel about the outparty remains less clear. How one’s view of the outparty changes in reaction to interparty interaction surely has downstream consequences on other political outcomes of social interaction, such as if information was distorted, if participation was hampered, and more. Therefore, it is important to also shed light on the immediate social psychological outcomes of social interaction, such as how we feel and think about outparty members (e.g., Mutz 2002).

Additionally, it is important to study the consequences of social interaction on the feelings and perceptions of the outparty because we can derive two plausible expectations about this process from the literature. Consider the view of American partisanship offered by Mason (2018), who calls partisanship, now aligned with many other identities, a "mega-identity" which heightens feelings of anger, competition with the outparty, and a need to "win" not just in terms of political interests but in terms of protecting their partisan "team." On the one hand, it follows that these feelings of anger, competition, and threat might surface at the prospect of conversation with an opposing party member. Thus, the drive to maintain a win for one’s partisan team could lead to an interaction that fails to improve, or even worsens, one’s negative view of the outparty. On the other hand, social interaction with outparty members could offer an opportunity to improve negative outparty attitudes as a form of intergroup contact, which is considered "one of psychology’s most effective strategies for improving intergroup relations" (pg. 5 Dovidio, Gaertner and Kawakami 2003). While Allport’s

---

1For example, Better Angles (https://www.better-angels.org/), Village Square (https://tlh.villagesquare.us/), American Public Square (https://americanpublicsquare.org/), Listen First Project (http://www.listenfirstproject.org/), and many more.
formulation of the 'contact hypothesis' stated four conditions the contact situation must meet to reduce negative outgroup attitudes (Allport 1954), multiple meta-analyses have since shown that contact, regardless of if it meets these conditions, has the tendency to improve negative outgroup evaluations (Paluck, Green and Green 2019; Pettigrew and Tropp 2006). In sum, the political science literature lacks a clear understanding of how social interaction with an opposing party member alters, if at all, the way partisans feel and think about the outparty.

3 Consequences of social interaction on outparty affect and stereotypes

In this section, I propose a resolution to the implicit tension in the political science literature between social identity theory and theories of intergroup contact. Specifically, I outline how social identity theory is a useful basis for understanding the consequences of contact with an outparty member, and I derive three expectations: that (1) non-political conversation and (2) political conversation with an outparty member improves a partisans’ negative view of the outparty at large, but that (3) non-political conversation will be more effective than political conversation at doing so.

3.1 Partisanship as a social identity

Social identity theory (e.g., Tajfel and Turner 1979) and the related self-categorization theory (e.g., Turner et al. 1987) present one framework for understanding intergroup, including interparty, biases and prejudices (Green, Palmquist and Schickler 2004; Greene 1999; Iyengar, Sood and Lelkes 2012). According to this framework, individuals associate with groups as a cognitive tool to understand their place in a complex, social world. As a consequence of forming a social identity, an individual’s sense of self becomes bound to the group, so maintaining a positive sense of self is tied to maintaining a positive view of the ingroup. Because the ingroup is understood in comparison to the outgroup, social identity theory hypothesizes that individuals are motivated to positively differentiate the ingroup from the outgroup——I like "us" more than "them." Research supports this hypothesis, showing that partisans express explicit and implicit favoritism, or bias, for members of the inparty even in non-political settings (e.g., Iyengar and Westwood 2015).

When partisanship is salient in a social interaction, individuals initially categorize themselves
and others as inparty or outparty members. There are two main consequences of this. First, when an individual self-categorizes, they comprehend and act within the interaction in accordance with how they see themselves as "a Republican" or "a Democrat" (Turner et al. 1987). Second, when a partisan categorizes an outparty member as such, they depersonalize the outparty member, thus viewing the outparty member as an oversimplified prototype of the broader group (Hogg and Reid 2006; Tajfel 1981). Affective and cognitive responses toward the outparty as a whole are applied to the outparty member. As discussed above, partisans hold a negative representation of the outparty that takes the form of a negative affect and negative trait stereotypes. This negative representation of the outparty will be used to form expectations about the outparty member’s beliefs and behavior at the start of a social interaction.

3.2 Social interactions as contact

In light of what the social identity perspective says about an individual’s initial understanding of social interaction with an outparty member, how might interparty contact unfold? Can contact alter the initial, negative representation of an outgroup member? I expect that contact can help partisans move beyond a prototypical understanding of the outparty member when personal, individuating information is shared (e.g., Brewer and Miller 1988, 1984; Ensari et al. 2012; Miller 2002). When such information is shared, partisanship can shift from being the most useful, or even the only, dimension shaping an understanding of the outparty member (and one’s self). Instead of categorizing an outparty member as such, social interaction allows the outparty member to be understood better as an individual person, breaking down the initially relied upon inparty/outparty categories.

However, the benefits of contact are limited if they extend to only the outgroup member involved in the immediate social interaction. Can contact alter the negative representation of the outgroup as a whole? While individuals initially understand a social interaction by categorizing—"us" verses "them," social interaction undermines this categorization by fostering interpersonal comparisons. Future interactions with outparty members then benefit from the personalized, rather than prototypical, representations of the outparty formed from the interparty contact (Brewer and Miller 1988, 1984). Moreover, evidence suggests that personalized contact improves biased outgroup affect and perceptions by interfering with the inclination to make sense of future social interactions based on those ingroup/outgroup categories (e.g., Bettencourt et al. 1992; Miller, Brewer and Edwards 1985).
Therefore, I expect that interparty interactions will undermine the usefulness of partisan identity as a basis for making sense of future social interactions, and as a result, contact will improve a biased, oversimplified representation of the outparty.

However, different types of social interaction may lead to variation in how effective contact is at fostering a personal, rather than prototypical, understanding of the outparty member. Consider what partisans talk about. It is relatively easy to imagine non-political social interactions that feature the sharing of personal information—talking about family or hobbies. Even though conversation about family, hobbies, or even your pets at home can cue partisanship (Hetherington and Weiler 2018), non-political conversations encourage partisans to understand each other as individuals, beyond (potentially incorrect) stereotyped views, group associations, and traits (e.g., Ahler and Sood 2018; Iyengar, Sood and Lelkes 2012; Levendusky and Malhotra 2015). Therefore, I expect non-political conversations can improve generalized outgroup affect and decrease the reliance on negative outgroup stereotypes in future interactions.

**Non-political contact hypothesis**: Non-political social interaction as contact with an outparty member, relative to no contact at all, improves (1) negative outparty affect and (2) negative outparty stereotypes.

It is less clear if social interactions can improve the negative view of the outparty even when conversation is about politics. However, I expect that in the context of political conversation, both non-political and political information can be shared that personalizes the outgroup member. While the outparty member begins as just one undifferentiated member of the outgroup, this understanding can be replaced with an understanding of them as an individual person, on both non-political and political dimensions. Therefore, I expect political conversations can also improve generalized outgroup affect and decrease the reliance on negative outgroup stereotypes in future interactions.

**Political contact hypothesis**: Political social interaction as contact with an outparty member, relative to no contact at all, improves (1) negative outparty affect and (2) negative outparty stereotypes.

While I expect that both non-political and political conversation can facilitate the benefits of contact generalizing to the outparty as a whole, there are important differences between a non-political and political conversations to consider. Partisan group membership is inherently more salient when talking about politics, heightening the initial motivation to secure a win for
one’s partisan team. So when talking about politics, it may be more difficult to move beyond an understanding of the social interaction based on an inparty/outparty categorization. Therefore, I expect non-political conversation will be more effective than political conversation for improving negative outparty affect and stereotypes.

**Topic hypothesis**: Non-political conversation improves (1) negative outparty affect and (2) negative outparty stereotypes more than political conversation.

In summary, I expect that social interaction across party lines tends to improve a partisans’ negative view of the outparty at large. I expect that non-political and political conversations are both effective settings to personalize the outparty member, break down inparty/outparty categories, and improve negativity toward the outparty. However, because partisan identities become more salient when conversation turns to politics, I expect non-political conversation will be more effective than political conversation at improving how partisans feel and think about the outparty.

4 Experiment

To assess the consequences of social interaction on outparty affect and stereotypes, I conducted an experiment on Amazon Mechanical Turk (MTurk) involving social interaction across party lines. The experiment required the following four steps, outlined in Figure 1. First, a set of potential participants took a pre-treatment survey to gather relevant pre-treatment covariates. At the conclusion of the survey, participants were asked if they would be willing to return for a follow-up task involving an "online chat with another Worker or writing a short essay."

Second, I used the pre-treatment survey responses of participants willing to return for the follow-up task to randomize participants into partnerships, each containing one Republican and one Democrat. Then, I randomly assigned conversation partnerships to one of three experimental conditions: no contact with partner (instead write an individual short essay), (2) contact with partner and discuss a non-political topic, or (3) contact with partner and discuss a political topic. Participants selected for the experiment were invited via email to complete the follow-up task. Having participants return for the experiment at all, let alone at the same time, presented a difficult coordination task. To minimize attrition between the pre-treatment survey and returning for the experiment, participants took the pre-treatment survey 10-30 minutes before the experiment. With
Figure 1: Experimental stages

Note: Visualization of participants’ and researcher’s roles throughout the four stages of the experiment.

the remaining 10 minutes, I randomized participants into partnerships, assigned partnerships to experimental conditions, emailed chosen participants 5 minutes before the experiment was live, and emailed chosen participants once more when experiment was live.

Third, participants selected for the experiment returned to complete the follow-up task where they spent a minimum of eight minutes writing a short essay or conversing with their assigned partner. Fourth, after completing their assigned task, participants proceeded to a post-treatment survey to assess how their feelings and perceptions about the outparty may have changed.

In what follows, I discuss several of these steps’ details—the three experimental conditions, the measurement of outcome variables, how partnerships and treatment were randomly assigned via a blocked cluster experimental design, and finally, how social interaction occurred via an online chat app.

4.1 Experimental conditions

The short essay and conversation prompts are shown in Table 1. Specifically, for partners assigned to have no contact with their outparty partner, each individual wrote separately about the meaning of life. For those assigned to the non-political contact condition, participants talked with their outparty partner about the meaning of life. I selected this topic because previous research has investigated how to induce close relationships, and thus foster personalized contact, in a laboratory experiment setting, finding that participants grow closer during a short interaction when
Table 1: Instructions for experimental conditions

<table>
<thead>
<tr>
<th>No contact</th>
<th>Non-political contact</th>
<th>Political contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please share your thoughts on the meaning of life. A conversation partner will not be joining you. Rather, we ask that you write about the meaning of life independently by sending messages in the chat box below. For example, survey research shows that many people mention family as the most important sources of meaning in their life. Survey research also shows that other people mention career, money, faith, friends, and hobbies as the most important source of meaning in their life. What do you think?</td>
<td>We’ve randomly assigned you a partner that belongs to or leans toward the [Republican/Democratic] party. Please have a conversation with them about the meaning of life. Specifically, we are interested in you sharing what you think makes life meaningful and learning your conversation partner’s thoughts as someone that might hold different values and beliefs. For example, survey research shows that many people mention family as the most important sources of meaning in their life. Survey research also shows that other people mention career, money, faith, friends, and hobbies as the most important source of meaning in their life. What do you think?</td>
<td>We’ve randomly assigned you a partner that belongs to or leans toward the [Republican/Democratic] party. Please have a conversation with them about gun control. Specifically, we are interested in you sharing your opinion on gun control and learning your conversation partner’s opinion as someone that might hold different values and beliefs. For example, survey research shows that some people believe it is more important to protect the right of Americans to own guns than control gun ownership, while others believe the opposite. Survey research also shows that some people believe making it harder to own guns would result in fewer mass shootings, while others believe this would make no difference. What do you think?</td>
</tr>
</tbody>
</table>

Note: Instructions for three experimental conditions. References to survey research included in the prompts come from recent Pew Research Center polls (Pew Research Center 2017, 2018).

communicating about 'deep' (i.e., What is the meaning of life?) rather than 'shallow' questions (e.g., What is your name? Where are your from?) (Sedikides et al. 1999; Tu, Shaw and Fishbach 2015). For those assigned to the political contact condition, participants conversed with their outparty partner about gun control. I selected this topic because it is a political issue salient to the average American so most participants are likely to have opinions they can converse about for a few minutes. Importantly, since intergroup contact implies outgroup membership is known during the contact situation, participants are told the partisanship of their conversation partner.

4.2 Outcome measures

I have two outcomes of interest. The first uses feeling thermometer scales to measure affect toward the outparty. Specifically, I asked the standard 101-point feeling thermometer, where larger values indicate more favorable or "warm" feelings toward that person or group (e.g., Iyengar, Sood and
Respondents rate both "Republicans across the country" and "Democrats across the country" in pre- and post-treatment surveys. My first outcome of interest is how contact can alter generalized outparty affect, which I define as the difference between pre- and post-treatment outparty feeling thermometer ratings. The second set of outcomes assess how contact can alter perceptions of the outparty, which I measure by asking participants to rate, using a five point Likert scale, how well several traits describe members of each political party (e.g., Levendusky 2018). Trait ratings were asked in the post-treatment survey only.

Respondents also indicate their partisan identification in the pre-treatment survey. Participants who choose Independent or Other are asked toward which party they lean. Due to the evidence that "leaners" behave similarly as partisans (Greene 1999; Iyengar and Westwood 2015; Pew Research Center 2019), I collapse Independents who lean toward one party into that party. Using partisan identification, I identify each individual’s outparty in order to construct the outcome measures of outparty affect and perceptions.

4.3 Experimental design

Lab experiments involving social interaction amongst participants, like this one, are common across the political discussion and deliberation literature (e.g., Druckman and Nelson 2003; Karpowitz, Mendelberg and Shaker 2012; Klar 2014); however, several methodological and practical concerns arise with this type of experiment. Not only does social interaction complicate a researcher’s design and subsequent data analysis, but small sample sizes, imbalance across experimental conditions, and more have implications for efficiency of estimation and power of hypothesis tests. And as a practical matter, experimental studies involving participant interaction are resource-intensive, often

---

2 Feeling thermometer question wording is, "Please rate the following groups using the following thermometer. Ratings between 50 and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 and 50 degrees mean that you don’t feel favorable toward and don’t care too much for that group. You would rate a group at the 50 degree mark if you don’t feel particularly warm or cold toward the group."

3 When using feeling thermometers and trait rating, research shows it is important to specify the group you want participants to rate beyond "the Democratic party" or "Democrats," for example (Druckman and Levendusky 2019). I ask respondents to rate "Democrats across the country" to target not only Democratic voters, but the entire outgroup that identifies as a Democrat.

4 Partisan identification question wording is, "Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?" A follow up question asks strength of partisan identity, "Would you call yourself a strong [Republican/Democrat] or not a very strong [Republican/Democrat]?” Finally, participants indicating Independent or Other are asked, "Do you think of yourself as closer to the Republican or Democratic party?"

5 "Lab experiment," as used here, refers to a context in which researchers have a high degree of control, particularly in the randomization of participants into groups and the randomization of treatment and control assignment. Lab experiments, by this definition, are not necessarily conducted in an academic laboratory.
prohibitively so, largely requiring an academic lab and existing subject pool.\(^6\)

To rigorously test the hypotheses derived in Section 3, I sought to address several of these methodological and practical concerns through two specific approaches to the experiment. First, I implemented a blocked cluster experimental design to improve efficiency of my estimation, among other advantages. Second, I developed a chat software to more easily allow for participant social interaction. In what follows, I briefly discuss each of these approaches in turn.

### 4.3.1 Blocked cluster design

For this experiment, I chose a blocked cluster design because (1) randomly assigning treatment at the *cluster*-level (here, conversation-level) is appropriate due to inherent interference between participants within a social interaction, and (2) randomly assigning treatment within *blocks* of clusters improves efficiency in estimation (e.g., Moore 2012). However, a blocked cluster design is typically used for field experiments which feature pre-existing clusters, such as cities or classrooms (e.g., Imai et al. 2009). To implement this design for a lab experiment, the researcher must somehow assign units to clusters (i.e., individuals to groups or partnerships). While guidance and tools exist for blocking (e.g., Moore 2012) and blocking with pre-existing clusters (e.g., Imai et al. 2009), it is less clear how to simultaneously block and cluster units. Therefore, I created an algorithm to construct a blocked, randomized cluster design.\(^7\) Importantly, this design allows me to ensure random assignment of what clusters (i.e., conversation partnerships) form, to improve efficiency of my estimation by block randomizing the treatment (at the cluster-level), and more.\(^8\)

Figure 2 outlines the five steps of my blocked, randomized cluster design algorithm. I use the hypotheses from this paper as an example. There are a few specifics about this example to highlight before explaining the steps of the algorithm. First, for this example, I want three similar clusters per block because I have three experimental conditions. This algorithm is generalizable to any number of experimental manipulations. Second, I want each cluster to have two participants. This

---

\(^6\) Appendix A discusses these methodological and practical concerns in more detail.

\(^7\) It is important to note that the proposed blocked, randomized cluster experimental design is certainly more complicated than one featuring completely randomized assignment of clusters and of the treatment. If the costs, in terms of complication, outweigh the benefits, I would not recommend this design. For example, if a large-\(n\) study is possible, then completely randomized groups and treatment assignment may be sufficient. Assessing power, efficiency, and more via DeclareDesign will prove helpful in these design decisions (Blair et al. 2019; Blair and Fultz 2019).

\(^8\) Appendix C presents simulations results suggesting the proposed blocked, randomized cluster design improves efficiency and power over a design featuring simple randomization of individuals to partnerships and treatments assignments.
algorithm is generalizable to any number of units per cluster. Third, I want each cluster to feature
one Republican and one Democrat. For this reason, I’ll call partisanship my "clustering constraint," or the variable the created clusters will be constrained to reflect. This algorithm is generalizable to any clustering constraint, such as disagreeable attitudes on the topic of discussion, different gender, or there doesn’t need to be one at all. Importantly, the clustering constraint must apply to all clusters to ensure the benefits of balance achieved by blocking.\(^9\)

The first step of the algorithm, demonstrated in the first plot of Figure 2, is to identify relevant blocking covariates and the clustering constraint, if any. For simplicity, I plot participants on only two dimensions—education and age. Because these variables likely affect the extent to which participants will change their outparty affect, I block on these variables to control for this variation. In step 1, I also indicate if the participants are Republican or Democrat because this is the variable all created clusters will be constrained to reflect—every cluster will have one Republican and one Democrat.

The second step is to identify temporary groupings of \(n\) similar units with respect to the clustering constraint, where \(n\) is the number of experimental manipulations.\(^10\) The second plot in Figure 2 shows the groupings of three similar units, conditional on partisan identification. Importantly,

\(^9\)A researcher may want to consider group composition as a treatment, such as creating same-party groups as a control condition for opposite-party conversation. In this case, one could imagine randomly assigning units to clusters to reflect one of these two clustering constraints (i.e., experimental conditions). However, careful consideration would need to be made to ensure doing so does not induce imbalance across experimental conditions on other covariates likely to impact the outcome of interest. Additionally, considering group composition as treatment (rather than fixed, pre-treatment) complicates consideration of a unit’s potential outcomes. This type of design is beyond the scope of this paper.

\(^10\)I create these temporary groupings using the blockTools statistical software (Moore 2016) with the optimum greedy algorithm and the Mahalanobis distance metric (Moore 2012). Details on the specific variables used for this step are available in Appendix D.
these groupings are not the clusters; rather, they are temporary groupings of similar units used to facilitate the creation of blocked and randomized clusters in subsequent steps.

The third step finishes the process of creating the blocks. I randomly assign each temporary grouping to another temporary set of units, conditional on the clustering constraint, if any. For this example, one group of similar Democrats is randomly assigned to one group of similar Republicans. These six individuals represent one block. Any remaining groups or individuals not assigned to blocks (as shown in gray) are discarded. It may seem counterintuitive to finalize the blocks before finalizing the specific cluster assignments. However, creating the blocks first ensures that the randomly assigned clusters are as similar as possible within each block.

The fourth step is to randomly assign clusters. Within each block, I randomly assign one unit from each temporary grouping to a unit from the other grouping. The fourth plot of Figure 2 shows this process for one block—units within the Democrat grouping and Republican grouping are randomly assigned to each other. The result is three similar, randomly assigned clusters grouped together in a block. Finally, with the created blocks and clusters in hand, treatment is randomly assigned at the cluster-level within each block as in any blocked cluster design. The fifth plot of Figure 2 demonstrates this step.

It is important to stress two features of this algorithm: cluster-level difference is minimized within and randomized across blocks! For example, Figure 2 shows that the clusters within this block each feature a young, less educated Democrat and an older, more educated Republican. The goal of blocking is to achieve this similarity at the cluster-level within a block—step 2 of Figure 2 is responsible for minimizing cluster-level differences in the finalized blocks. On the other hand, I do not want all of the conversation partnerships to be between very dissimilar Republicans and Democrats. How different the Republican and Democrat pairings are across blocks varies—step 3 of Figure 2 is responsible for randomizing cluster-level differences across blocks. While the clusters in the block in Figure 2 feature partners who are all very different in the same ways, another block may feature partners who are all very similar.

### 4.3.2 Chatter conversation software

In addition to addressing several methodological concerns via the blocked, randomized cluster experimental design, I sought to overcome practical concerns that arise when conducting experiments...
involving social interaction. To do so, I took a novel approach to how the social interaction amongst participants would occur. I designed a software called "Chatter" so participants can have real-time, written conversations online. Full details on Chatter are available in Appendix B. Chatter allows me to emulate a real social experience for the study participants without an existing laboratory or a participant pool. Moreover, Chatter provides me full experimental control—participants are filtered into chatrooms with the partners and treatments pre-assigned via the blocked, randomized cluster design.

5 Data and results

I fielded the experiment between August 15-22, 2019 and January 13-15, 2020 on MTurk with eight separate rounds of data collection. In total, 1,632 unique MTurk Workers took the pre-treatment survey and a subset of 630 were selected via the blocked, randomized cluster experimental design algorithm. Therefore, the design included 630 participants, formed into 315 partnerships/clusters, and nested within 105 blocks.

Despite the quick timeline and reminders outlined above, chosen participants did attrit between the pre-treatment survey and returning for the experiment. And, a single participant’s attrition impacts their cluster and block. The sample contains 238 (75.5%) full partnerships and 44 (41.9%) full blocks.\footnote{Appendix E shows details on attrition by round of data collection.} In what follows, all blocks for a which any participant did not follow up are dropped.\footnote{Appendix I shows results for all full partnerships/clusters which are consistent with results when including only full blocks.} This is an advantage of block randomized designs. Dropping entire blocks allows for the loss of clusters without losing balance on blocked covariates across experimental conditions, which can generate bias (e.g., King et al. 2007). After doing so, the sample used in subsequent analyses contains a total of 264 participants: 44 pairs in the control condition, 44 pairs in the non-political contact condition, and 44 pairs in the political contact condition. Importantly, no participants attrited post-treatment, which could bias results if participants attrited as a function of treatment assignment, such as after seeing they were assigned to talk politics.
5.1 Manipulation check

Before considering the effects of social interaction on outparty affect and stereotypes, I present results suggesting the participants took the exercise seriously and engaged in thorough, on-topic conversation.\(^{13}\) First, it is evident the participants engage in the assigned exercise of either a short essay or conversation. A median number of 17.5 messages were sent across non-political conversation and 14 messages across political conversations. The median number of words exchanged were 222, 269, and 299 in the no contact, non-political contact, and political contact conditions, respectively. These summary statistics demonstrate the participants engaged in the exercise. Moreover, participants discussed their assigned topic. The word "gun" appeared 386 times across the conversations assigned to talk about gun control, and not at all in the other experimental conditions. The word "meaning" appeared 394 times in the control condition, 119 times in the non-political contact condition, and not at all in the political contact condition.\(^ {14}\)

5.2 Randomization inference hypothesis tests

Given the evidence that the participants engaged in their assigned type of social interaction, I next assess my claims regarding the consequences of social interaction with an outparty member.

I test my hypotheses using randomization inference (Fisher 1935). Because my sample size is relatively small, and as I demonstrate below, my outcomes are not distributed normally, I use randomization inference to avoid appealing to large sample approximations or modeling assumptions. Additionally, randomization inference is straightforward given the blocked cluster randomization procedure I use for assigning conversation partnerships to experimental conditions. Rather than state a null hypothesis of no average treatment effect, randomization inference depends on a sharp null hypothesis of no effect for all units. Under this hypothesis, I can then state all potential outcomes for all units—given the sharp null hypothesis of no effect, the potential outcome for a unit under the control condition is no different than their outcome under treatment. Then, since I know the exact randomization procedure, it is possible to list all permutations of treatment assignment and obtain all hypothetical results of the experiment (using some test statistic) under the sharp null hypothesis.

\(^{13}\)Appendix F provides an example from each of the three experimental conditions of on-topic, active participation in the exercise.

\(^{14}\)Appendix G presents consistent results across additional words and phrases.
Figure 3: Density of outcome variable, change in outparty affect, by experimental condition

Note: Density of outcome variable, change in outparty affect, by experimental condition considering full blocks. Outcome measured as change in outparty affect measured using pre- and post-treatment ratings of the outparty on the 101-point feeling thermometer. Grey density shows outcome for the no contact condition, blue density shows outcome for the political contact condition, and red density shows outcome for the non-political contact condition. Dashed lines display the mean change for each condition. Change for the no contact condition is tightly concentrated around 0. Change for each contact condition is skewed positively, indicating a general improvement in outparty affect.

null. Given this sampling distribution of the test statistic under the sharp null, I calculate $p$-values to judge if the result of my experiment is unusual (i.e., significant).

I use the suite of tools available in the DeclareDesign software to declare my blocked cluster randomization procedure (Blair et al. 2019), and I use the ri2 software (Coppock 2019) to conduct the randomization inference. Because there are $2^{44}$ possible random assignments for the following tests, in what follows, I find approximate $p$-values using a random sample of 50,000 random assignments consistent with my design.

I first assess contact’s effects on outparty affect before turning to contact’s effect on outparty trait stereotypes. Figure 3 shows a summary of the change in outparty affect outcome variable. Change is tightly concentrated around zero for the no contact condition, while each of the contact conditions shows considerable positive change in affect. Improvement in outparty affect is notably similar across non-political and political conversations.

I use randomization inference to test my three hypotheses regarding the benefits of contact with an outparty member on outparty affect. Note I consider this a three-armed experiment, so I assess pairwise comparisons between each experimental condition.\footnote{Out of concerns of dropping too many blocks due to attrition, I omitted a no contact & political topic condition. After gaining more experience with the MTurk Workers and scaling the Chatter software, I plan to conduct the full test.} Table 8 reports the observed average
Table 2: Significance tests of average treatment effect estimates, outcome change in outparty affect

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-political contact vs. no contact</td>
<td>9.17</td>
<td>≈0</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>7.47</td>
<td>.002</td>
</tr>
<tr>
<td>Non-political vs. political contact</td>
<td>1.70</td>
<td>.58</td>
</tr>
</tbody>
</table>

Note: Results considering full blocks. Difference-in-differences estimates of average treatment effect and randomization inference p-values under the sharp null hypothesis of no treatment effect. Dependent variable is individual-level change in outparty affect. Two-sided p-values are reported. Both non-political contact and political contact improved outparty affect relative to no contact.

treatment effect estimate and the p-value associated with a test of this estimate against the sharp null hypothesis of no effect. I use the difference-in-difference estimator for the average treatment effect because my outcome is measuring "change scores"—change in outparty affect from pre-test to post-test. The first row in Table 8 reports results for my first hypothesis—that non-political interparty contact improves negative outparty affect. I find that non-political contact improved outparty affect on average by 9.2 degrees, relative to no contact. Moreover, the second row in Table 8 reports results for my second hypothesis—that political interparty contact improves negative outparty affect. I also find support for this hypothesis as political contact improved outparty affect on average by 7.4 degrees, relative to no contact. Finally, I assess my third hypothesis—that non-political conversation would be more effective than political conversation at improving outparty affect. Surprisingly, I do not find evidence of non-political conversation being more effective political conversation at improving intergroup attitudes.

To help put these results into context, I summarize participants’ changed outparty feeling thermometer ratings relative to meaningful points on the scale (e.g., Levendusky 2018). First, consider the percentage of participants rating the outparty "warmly," or greater than or equal to 50 on feeling thermometer, after contact—44% of those in non-political condition and 38% of those in the political condition rated the outparty in this way while only 25% of participants who did not experience interparty contact rated the outparty favorably. Additionally, at the "cold" end of the feeling thermometer, consider a very unfavorable rating of less than or equal to 5. Only 17% of those in the non-political condition and 15% of those in political condition rated outparty in this way, while 28% of those who did not experience outparty contact rated the outparty with such an
Table 3: Significance tests of average treatment effect estimates, outcome perceptions of the outparty

<table>
<thead>
<tr>
<th></th>
<th>'Openminded'</th>
<th></th>
<th>'Mean'</th>
<th></th>
<th>'Hypocritical'</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>p-value</td>
<td>Estimate</td>
<td>p-value</td>
<td>Estimate</td>
<td>p-value</td>
</tr>
<tr>
<td>Non-political contact vs. no contact</td>
<td>.39</td>
<td>.020</td>
<td>-.59</td>
<td>≈0</td>
<td>-.59</td>
<td>≈0</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>.35</td>
<td>.053</td>
<td>-.66</td>
<td>≈0</td>
<td>-.56</td>
<td>≈0</td>
</tr>
<tr>
<td>Non-political contact vs. political contact</td>
<td>.034</td>
<td>.89</td>
<td>.068</td>
<td>.737</td>
<td>-.034</td>
<td>.878</td>
</tr>
</tbody>
</table>

Note: Difference-in-means estimates of average treatment effect and randomization inference p-values under the sharp null hypothesis of no treatment effect. Dependent variable is agreement, on a five point scale with higher values indicating more agreement, with how well the trait describes supporters of the outparty. Two-sided p-values are reported. Results for additional traits are reported in Appendix H with largely consistent results.

extremely unfavorable rating.

I’ve provided evidence that social interaction can alter how partisans feel about the outparty, and I now turn to assess if contact can alter how partisans think about, or perceive, outparty members. The outcomes of interest are respondents’ level of agreement (on a five point scale, higher values indicating more agreement) with how well several traits describe supporters of the outparty. I use the difference-in-means estimator to estimate average treatment effects. Table 5 reports these estimates and randomization inference p-values under the sharp null of no effect for three traits: openminded, mean, and hypocritical. When respondents had interparty social interaction, whether non-political or political, they were more inclined to ascribe positive traits to the outparty (i.e., openminded) and less inclined to ascribe negative traits to the outparty (i.e., mean or hypocritical). The final row of Table 5 reports results regarding the difference between non-political and political contact. I again do not find any evidence in support of my hypothesis that non-political contact would be more effective than political contact at improving a partisan’s negative view of the outparty. Surprisingly, political and non-political conversations do not have a distinguishable difference in their effect on outparty stereotypes nor outparty affect.

6 Conclusion

Despite the ample evidence documenting a macro-level trend of increasing negativity toward members of the opposite political party, relatively less research has investigated the consequences of micro-level social interactions with outparty members on this trend. This paper addresses this question with an

---

16 Appendix H reports largely consistent results for the additional traits surveyed.
experiment involving actual social interaction among partisans, and shows that social interaction can result in a sizable increase in outparty affect and a disinclination to describe outparty members by negative stereotypes. In doing so, this research address a concern in the literature to appropriately characterize the extent to which Americans are "affectively polarized"—increasingly liking their inparty and disliking the outparty—and what consequences this type of polarization has. While this research characterizes Americans’ views of outparty as biased, overgeneralized, and potentially inaccurate, it also illuminates the limits of our partisan identities by showing social interaction has the power to break down inparty/outparty categories and correct for heightened outparty negativity.

While I’ve found that social interactions can improve how partisans feel about and think about outparty members, a question left for future research is when social interactions improve affect and perceptions outside of the environment constructed for this research. For example, this experiment featured online conversation, limiting the external validity of these findings as applied to in-person interactions where physical appearance and body language are additional guides to social interaction that may influence how it unfolds and what effects it has on subsequent outcomes.

Moreover, this research does not reflect the role self-selection plays in this process of social interaction, as partisans may try to choose when and when not to engage in social interactions with outparty members. Research shows that anticipating political discussion makes people anxious (Carlson and Settle 2016). It follows that people prefer to avoid political discussion, especially when it is disagreeable (Gerber et al. 2012) or with an outparty member (Settle and Carlson 2019).

But, while research finds some people prefer avoid political conversation or social interaction with outparty members, sometimes these interactions occur beyond our control. Walsh contends that "Much political interaction occurs not among people who make a point to specifically talk about politics but emerges instead from the social processes of people chatting with one another" (pg. 35 Cramer Walsh 2004). Carlson and Settle echo this view, stating "Regardless of their interest in it, people experience politics through offhand observations and interactions with people in their daily lives" (pg. 819 Carlson and Settle 2016). Moreover, a recent large-scale, full-network study supports the idea that talking politics is more an incidental than it is a purposive exercise (Minozzi et al. 2019). If we take the incidental model of political discussion seriously, then talking politics is often unanticipated, it is hard to avoid altogether, and everyone is subject to experiencing some political talk in their daily lives, and this research begins illuminating the consequences of it.
Another limitation of this experimental design is that it involved only two individuals, one from each party. While this helps satisfy one of Allport’s conditions for contact to improve outgroup prejudice—equal status in the contact situation—not all social interactions will avoid having a majority opinion apparent in the interaction. This is an important consideration because research shows that when politics arises in a discussion, people tend shield their own views and conform to majority opinion (Carlson and Settle 2016). Related to conformity within social interactions, this research does not account for the role social sanctioning may play in political and non-political interactions that occur in Americans’ everyday lives. It is left for future research to speak to how different group compositions and pre-existing relationships may impact the effectiveness of social interaction as a strategy for combating negative intergroup attitudes.

Finally, this research theorizes about and studies positive contact. However, research shows that a single negative instance of contact has a stronger, negative effect on affect than does any single instance of positive contact (Barlow et al. 2012). Therefore, future research should consider under what conditions partisan social interactions go awry, such as when a partisan feels like their group or "team" is being threatened. While this research studied the topic of gun control, electoral politics or issues appealing to more deeply held values may activate stronger emotions and lead partisans to use social interactions to fortify the walls of their inparty/outparty categories.
References


Grumet, Jason. 2019. “Why American families must talk about politics this Thanksgiving.” *The


Appendix A  Methodological and practical concerns in the experimental study of social interaction

Political scientists often use experimental research in which participants have actual social interaction to test their hypotheses (e.g., Druckman and Nelson 2003; Karpowitz, Mendelberg and Shaker 2012; Klar 2014). Lab experiments across this broad range of research share a set of methodological and practical concerns. In this section, I outline several of these concerns after first walking through the setup of a typical lab experiment in this field.

The typical design for a lab experiment involving social interaction is the following. The researcher has a list of participants, maybe the students in their department’s Introduction to American Politics class. If the experiment requires small group discussion of four participants per group, the researcher randomly assigns participants into groups of four. Or, the experiment might involve a certain group composition, such as discussion amongst opposing party members. If so, the groups are randomly assigned with respect to this constraint. After random group assignment, treatment is assigned at the group-level. For example, the treatment might be whether or not the group engages in discussion (e.g., Klar 2014) or the decision rule governing the group’s decision-making (e.g., Karpowitz, Mendelberg and Shaker 2012). Finally, random assignment is typically used to assign groups to experimental conditions.

I discuss four main, interrelated methodological concerns that arise with experiments involving social interaction: sample size, balance across experimental conditions, and unit interference, and attrition. First, these studies often rely on relatively small sample sizes (e.g., 277 in Klar (2014) or 261 in Druckman and Nelson (2003)). Small sample sizes impacts the power of tests, decreasing the probability a researcher can reject the null hypothesis of no treatment effect. Additionally, while simple random assignment or complete random assignment of the treatment are straightforward, they can introduce inefficiencies, especially when coupled with small sample sizes. Because the sample size for these studies is usually small, a researcher could end up with an unlucky, unbalanced randomization. What’s more, balance across experimental conditions is not just relevant at the individual level, but it is relevant at the group level, as well. For example, for an experiment involving two individuals engaging in conversation, it may be problematic if all strong Republicans were, by chance, assigned to the same experimental condition. Additionally, it may be problematic if all groups who are ideologically-similar were, by chance, assigned to the same experimental condition. Block randomizing treatment assignment can ensured balanced randomization and improve inefficiencies (Moore 2012). Relatedly, this class of experiments inherently involves social influence, and therefore interference or spillover, within the assigned groups.

Beyond the study of conversation, lab experiments in political science featuring constructed interaction amongst participants take the form of trust and ultimatum games, for example (e.g., Alford and Hibbing 2007; Carlin and Love 2013).

Lab experiments involving participant discussion require a great deal of effort and coordination by the researcher and participants. Alternatively, to increase the number of completed discussion groups in the experiment, the researcher may construct groups based on which participants are available and present at any given time, thus without randomization.
Treatment, therefore, ought to be applied at the group level, introducing an added complexity to the experimental design and estimation of treatment effects. Randomizing groups, rather than individuals, also reduces efficiency in this class of experiments (Cornfield 1978).

As a practical matter, studies involving participant interaction are resource-intensive, often prohibitively so, largely requiring an academic lab and existing subject pool. Even with these resources in place, it is difficult to then coordinate participants into complete discussion groups. Therefore resources for hosting social interaction and coordinating the participants, all the while maintaining sample size, are practical concerns of experiments involving social interaction.

Appendix B  Chatter conversation software

As a practical matter, in order to study interparty conversations, I designed a software called 'Chatter' by which participants can have real-time, written conversations online. Chatter allows 2-10 participants to have a conversation via an interface similar to many messaging apps like iMessage or WhatsApp.

Figure 4 shows an example of Chatter’s user interface. Participants see instructions for the conversation above a box containing previously sent messages. A participant’s own messages appear unlabeled on the right in blue and other conversation members’ messages appear, labeled with a unique identifier, on the left. Participants also see a timer that counts down the time remaining and a 'Done' button which illuminates and activates when time expires.

Several features of Chatter facilitate experimentation involving conversations. First, Chatter allows the researcher to set up chatrooms so that the partnerships or groups created using the blocked, randomized cluster design can engage in social interaction. Second, Chatter allows the researcher to customize the conversational instructions shown to each participant. Specifically, in the experiment I explain below, treatment takes the form of the topic at the cluster-level, so conversation partners see the same instructions. But, each participant’s instructions are customized to include the partisan identity of their partner. Third, the researcher controls all additional chatroom and user settings, including what usernames are displayed (e.g., random sequence of letters, the same name to control perception of gender, etc.) and how long the conversation should last. Lastly, I use Chatter coupled with the Qualtrics survey software for survey administration and Amazon Mechanical Turk to recruit participants. However, experimentation using Chatter is generalizable to other survey platforms and other participant pools.

Chatter allows researchers to relatively easily emulate a real social experience without an existing laboratory or a participant pool. Moreover, as a software for online conversations, Chatter allows for diversity in the participant pool that is hard to come by when using in-person conversations, usually with students, faculty, and staff that are affiliated with a college campus (e.g., Karpowitz and Mendelberg 2014; Klar 2014) or that can be recruited within a few cities (e.g. Druckman, Levendusky and McLain 2017). A final practical advantage of Chatter is that a researcher can quickly conduct

---

19Chatter is a Ruby on Rails application backed by a Postgres database, deployed on Heroku.
Note: Chatter user interface. Instructions appear at the top of the page. Akin to other messaging software, an individual’s own messages appear on the right. Other users’ messages appear on the left. When the timer indicates no time is left, the "Done" button activates and redirects users to a post-conversation survey when clicked.

a large-n study involving conversations. Chatter allows for hundreds of conversations to happen simultaneously, which is difficult to achieve in the setting of an academic laboratory.

Appendix C Simulation study for proposed blocked, randomized cluster design

I ran a simple simulation to demonstrate the benefits of the blocked, randomized cluster design as opposed to what I will call a 'naive' design. The simulation mimics the construction of clusters used in this paper, where the goal is one Republican and one Democrat per cluster. The naive design creates clusters by (1) randomly choosing a participant who does not already have a partner assigned from the naive design and (2) randomly choosing a partner of the opposite partisan identification from the remaining un-partnered participants. Then, using simple random assignment, the naive
Note: Diagnosands of interest for the proposed blocked, randomized cluster design and the naive design, while varying sample size. Root mean squared error (RMSE) is reported in the first plot, mean absolute error (MAE) is reported in the second plot, and power of randomization inference hypothesis tests is reported in the third plot. Note that the proposed design is more efficient (in terms of RMSE and MAE) and has better power than the naive design.

For this simple example, the population is of size $N = 500$, each individual is assigned a partisan identification label, and three variables impact the outcome: $X_1 \sim N(0,1)$, $X_2 \sim Unif(0,1)$, and $X_3 \sim \chi^2_2$. There’s both individual-level error $u_i \sim N(0, 1)$ and cluster-level error $u_c \sim N(0,1)$.

Potential outcomes are a function of these variables: $Y_i(Z = 0) = X_1 + 2 \times X_2 + 3 \times X_3 + u_i + u_c$ and $Y_i(Z = 1) = X_1 + 2 \times X_2 + 3 \times X_3 + 1 + u_i + u_c$. My estimand is the average treatment effect. Importantly, I vary the number of clusters I sample from the population of fixed clusters from 25 to 150 clusters for the naive design. To vary the sample size in a comparable way for the blocked, randomized cluster design, I sample half as many fixed blocks. I use the difference in means estimate, and calculate $p$-values for hypothesis tests using randomization inference where I can specify the exact randomization procedure used for the blocked, randomized cluster design or the naive design.

I use the DeclareDesign and ri2 R packages (Blair and Fultz 2019; Coppock 2019) to conduct this simulation study. I conduct 500 simulations of the design for each sample size. I assess the performance of the two designs with three diagnosands: root mean squared error (RMSE), mean absolute error (MAE), and power with $\alpha = .05$. Figure 5 shows the results. The proposed blocked, cluster design improves efficiency by having a lower RMSE and MAE across the entire range of sample sizes, and particularly for smaller sample sizes. Additionally, the proposed experimental design rapidly improves the power of randomization inference hypothesis tests in this simulation as the sample size increases.
Appendix D  Details on blocked cluster design for experiment

The following provides more specific details for the block, randomized cluster algorithm used in the experiment, particularly for the blocking variables.

- Create trios of similar individuals within the same party
  - Using the `blockTools::block()` statistical software (Moore 2016) with the optimum greedy algorithm and the Mahalanobis distance metric (Moore 2012). Note is step does not create the blocks in full, despite the use of the `blockTools` statistical software. See Section 4.3.
  - Blocking variables are age and education (considered continuous); indicators for gender* and ethnicity, strength of partisan identification*; pre-treatment thermometer ratings of inparty, outparty, and President Trump*; a 6-item battery on pre-treatment gun views (considered continuous); a single pre-treatment item asking overall view regarding gun control*; personality trait estimates including 4 item adaptive versions of each Big 5 Trait (openness*, conscientiousness, agreeableness, neuroticism, and extraversion) (Costa and McCrae 2008), Systemizing and Empathizing Quotients (Baron-Cohen et al. 2003); and latitude and longitude. Variables marked with an * are up-weighted to have twice the weight of the other variables.
  - Block within subgroups of partisan identification. Those who indicated "Independent" or "other" for partisan identification are collapsed into the respective party toward which they lean.

- Shuffle individuals within trios.
- Simultaneously create blocks and interparty pairs by randomly assigning one Democratic trio to each Republican trio.
- Within each block of interparty pairs, randomly assign treatment at interparty pair-level.

Appendix E  Attrition details

<table>
<thead>
<tr>
<th>Round</th>
<th>Pre-test Participants (N)</th>
<th>Invited to Return (N)</th>
<th>Full Pairs (%)</th>
<th>Full Blocks (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>124</td>
<td>42</td>
<td>85.7%</td>
<td>71.4%</td>
</tr>
<tr>
<td>2</td>
<td>138</td>
<td>30</td>
<td>73.3%</td>
<td>40.0%</td>
</tr>
<tr>
<td>3</td>
<td>146</td>
<td>54</td>
<td>63.0%</td>
<td>22.2%</td>
</tr>
<tr>
<td>4</td>
<td>174</td>
<td>78</td>
<td>71.8%</td>
<td>30.8%</td>
</tr>
<tr>
<td>5</td>
<td>249</td>
<td>84</td>
<td>76.2%</td>
<td>42.9%</td>
</tr>
<tr>
<td>6</td>
<td>270</td>
<td>102</td>
<td>74.5%</td>
<td>35.3%</td>
</tr>
<tr>
<td>7</td>
<td>311</td>
<td>132</td>
<td>83.3%</td>
<td>59.1%</td>
</tr>
<tr>
<td>8</td>
<td>220</td>
<td>108</td>
<td>72.2%</td>
<td>33.3%</td>
</tr>
<tr>
<td>Overall</td>
<td>1632</td>
<td>630</td>
<td>75.6%</td>
<td>41.9%</td>
</tr>
</tbody>
</table>
Appendix F   Example short essays and conversations
I think that feeling content in where you are is really one of the deepest meanings in life. Having a beautiful connection with someone you love makes life feel so much more worth it. Being able to connect spiritually and find your inner peace is a great goal. Finding someone who is your soul match and a willing partner is the ultimate meaning of life. Not letting money or materialistic things rule you, but instead letting those things be a byproduct of what you love. Finding your true family, whether blood or not is a great goal in life. Realizing that sometimes you have to let go and let God is a great goal. Making sure your heart and your spirit are in a place of love is incredibly important in life. Not being around anyone who takes away your joy and your light is an important goal in life. All of these things together are pertinent to reaching your full potential aka the meaning of your life. There is no right and wrong answer to what the exact meaning is, but you have to look within yourself to see what your soul and spirit deem important. The true meaning of life is to feel fully content and at peace with your mind, body and spirit. The true meaning of life is love.

For me life is happiness and love. Happiness is when you are with the people you cared for and love. I think life will be dull or boring if we don’t have those people that we will cherish the most. They are the ones who makes the ride worthwhile. Life also is giving to others. It is sharing what you have and not asking in return. Life is when you cry and laugh the most. Life is when you are hurt but prefer to stand up. Life is still fighting despite the hardship. Life is loving all the positive things. Life is protecting our mother nature. Life is everything.

What do you think about the meaning of life? I think that the meaning of life is very simply to be good people and to spread love and positivity to others. I think that the mean of life is to enjoy every second we live on this earth. I think it’s important to notice the beauty in the simple things. I definitely agree with that, especially because we have no idea when our life is going to end and if there is something after this life. We definitely don’t appreciate the small things in life, we tend to take them for granted in my opinion. I know I sometimes blow things way out of proportion with regards to things going wrong in my life. But in reality, my problems really aren’t that big of a deal. I think that people are so focused on the long term they don’t think about the here and now. I’ll admit I’m one of them. I agree. I mean it’s good to focus on the future and prepare yourself but I think it’s also important to balance the future with the present as well. I see so many people setting themselves up for future success but in the present they’re miserable, working a ton of hours at their jobs and hating their lives because they have no free time. I think that people are so focused on the long term they don’t think about the here and now. I know I sometimes blow things way out of proportion with regards to things going wrong in my life. But in reality, my problems really aren’t that big of a deal. Yes I agree. I think that people are so focused on the long term they don’t think about the here and now.

For me life is happiness and love. Happiness is when you are with the people you cared for and love. I think life will be dull or boring if we don’t have those people that we will cherish the most. They are the ones who makes the ride worthwhile. Life also is giving to others. It is sharing what you have and not asking in return. Life is when you cry and laugh the most. Life is when you are hurt but prefer to stand up. Life is still fighting despite the hardship. Life is loving all the positive things. Life is protecting our mother nature. Life is everything.

What do you think about the meaning of life? I think that the meaning of life is very simply to be good people and to spread love and positivity to others. I think that the mean of life is to enjoy every second we live on this earth. I think it’s important to notice the beauty in the simple things. I definitely agree with that, especially because we have no idea when our life is going to end and if there is something after this life. We definitely don’t appreciate the small things in life, we tend to take them for granted in my opinion. I know I sometimes blow things way out of proportion with regards to things going wrong in my life. But in reality, my problems really aren’t that big of a deal. I think that people are so focused on the long term they don’t think about the here and now. I’ll admit I’m one of them. I agree. I mean it’s good to focus on the future and prepare yourself but I think it’s also important to balance the future with the present as well. I see so many people setting themselves up for future success but in the present they’re miserable, working a ton of hours at their jobs and hating their lives because they have no free time. I think that people are so focused on the long term they don’t think about the here and now. I know I sometimes blow things way out of proportion with regards to things going wrong in my life. But in reality, my problems really aren’t that big of a deal. Yes I agree. I think that people are so focused on the long term they don’t think about the here and now.

Table 4: Example short essays, conversations for experimental conditions

<table>
<thead>
<tr>
<th>No contact (control)</th>
<th>Non-political contact</th>
<th>Political contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1 I think that feeling content in where you are is really one of the deepest meanings in life. Having a beautiful connection with someone you love makes life feel so much more worth it. Being able to connect spiritually and find your inner peace is a great goal. Finding someone who is your soul match and a willing partner is the ultimate meaning of life. Not letting money or materialistic things rule you, but instead letting those things be a byproduct of what you love. Finding your true family, whether blood or not is a great goal in life. Realizing that sometimes you have to let go and let God is a great goal. Making sure your heart and your spirit are in a place of love is incredibly important in life. Not being around anyone who takes away your joy and your light is an important goal in life. All of these things together are pertinent to reaching your full potential aka the meaning of your life. There is no right and wrong answer to what the exact meaning is, but you have to look within yourself to see what your soul and spirit deem important. The true meaning of life is to feel fully content and at peace with your mind, body and spirit. The true meaning of life is love.</td>
<td>u3 I agree. I mean it’s good to focus on the future and prepare yourself but I think it’s also important to balance the future with the present as well. I see so many people setting themselves up for future success but in the present they’re miserable, working a ton of hours at their jobs and hating their lives because they have no free time. Exactly. That’s why I decided to travel and make the most of my youth. What I try to do is enjoy every single day. I try to find minor things that make me happy such as going out to a restaurant or playing a fun game on my phone.</td>
<td>u5 Hi, I am ready when you are. thank you. u6 I agree. I mean it’s good to focus on the future and prepare yourself but I think it’s also important to balance the future with the present as well. I see so many people setting themselves up for future success but in the present they’re miserable, working a ton of hours at their jobs and hating their lives because they have no free time. Exactly. That’s why I decided to travel and make the most of my youth. What I try to do is enjoy every single day. I try to find minor things that make me happy such as going out to a restaurant or playing a fun game on my phone.</td>
</tr>
<tr>
<td>u2 For me life is happiness and love. Happiness is when you are with the people you cared for and love. I think life will be dull or boring if we don’t have those people that we will cherish the most. They are the ones who makes the ride worthwhile. Life also is giving to others. It is sharing what you have and not asking in return. Life is when you cry and laugh the most. Life is when you are hurt but prefer to stand up. Life is still fighting despite the hardship. Life is loving all the positive things. Life is protecting our mother nature. Life is everything.</td>
<td>u4 I think it’s important to notice the beauty in the simple things. I definitely agree with that, especially because we have no idea when our life is going to end and if there is something after this life. We definitely don’t appreciate the small things in life, we tend to take them for granted in my opinion. I know I sometimes blow things way out of proportion with regards to things going wrong in my life. But in reality, my problems really aren’t that big of a deal. Many people worldwide are way less off than me which is why I try to appreciate everything that I have. It’s difficult at times though, sometimes little things add up and can stress you out and you forget to put that type of stuff into perspective. Yes I agree. I think that people are so focused on the long term they don’t think about the here and now.</td>
<td>u5 My thoughts on gun control are that there are plenty of laws we just need to follow them although maybe some gun types just aren’t needed in the general public. u6 My thoughts on gun control are that there are plenty of laws we just need to follow them although maybe some gun types just aren’t needed in the general public.</td>
</tr>
<tr>
<td></td>
<td>u5 My thoughts on gun control are that there are plenty of laws we just need to follow them although maybe some gun types just aren’t needed in the general public. u6 My thoughts on gun control are that there are plenty of laws we just need to follow them although maybe some gun types just aren’t needed in the general public.</td>
<td>u7 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
</tr>
<tr>
<td></td>
<td>u7 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
<td>u8 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
</tr>
<tr>
<td></td>
<td>u8 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
<td>u9 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
</tr>
<tr>
<td></td>
<td>u9 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
<td>u10 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
</tr>
<tr>
<td></td>
<td>u10 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
<td>u11 I know gun enthusiasts that like their assault rifles very much. I don’t know how we are going to keep guns out of the hands of the mentally ill....that is a grey area to regulate.</td>
</tr>
</tbody>
</table>
Appendix G  Manipulation check

The following table details how many times the following terms appear across all short essays or conversations for each experimental condition. Phrases such as "family," "faith," and "happiness" occur often, as expected, when talking about the meaning of life, and do not occur when participants are asked to discuss gun control. Likewise, "gun," "shooting," and "background check" occur often when discussing gun control but not when participants were asked to discuss the meaning of life.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>gun</th>
<th>shooting</th>
<th>background check</th>
<th>family</th>
<th>faith</th>
<th>meaning</th>
<th>happiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>No contact</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>111</td>
<td>14</td>
<td>394</td>
<td>42</td>
</tr>
<tr>
<td>Non-political contact</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>7</td>
<td>119</td>
<td>18</td>
</tr>
<tr>
<td>Political contact</td>
<td>386</td>
<td>38</td>
<td>38</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Appendix H  Additional outparty trait stereotypes

Table 5: Significance tests of average treatment effect estimates of outparty stereotypes

<table>
<thead>
<tr>
<th></th>
<th>&quot;Patriotic&quot;</th>
<th>&quot;Closeminded&quot;</th>
<th>&quot;Intelligent&quot;</th>
<th>&quot;Selfish&quot;</th>
<th>&quot;Honest&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Estimate</td>
<td>Estimate</td>
<td>Estimate</td>
<td>Estimate</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>P-value</td>
<td>P-value</td>
<td>P-value</td>
<td>P-value</td>
</tr>
<tr>
<td>Non-political contact vs. no contact</td>
<td>.125</td>
<td>.444</td>
<td>-.34</td>
<td>.021</td>
<td>.125</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>.341</td>
<td>.063</td>
<td>-.33</td>
<td>.046</td>
<td>.34</td>
</tr>
<tr>
<td>Non-political contact vs. political contact</td>
<td>-.216</td>
<td>.210</td>
<td>-.011</td>
<td>≈ 1</td>
<td>-.216</td>
</tr>
</tbody>
</table>

Note: Results considering full blocks. Randomization inference hypothesis tests with the difference in means test statistic under the sharp null hypothesis of no treatment effect. Dependent variable is agreement, on a five point scale with higher values indicating more agreement, with how well the trait describes supporters of the outparty. Two-sided p-values are reported.
Appendix I  Results including all full clusters

Results are consistent if I ignore the blocking element of the randomization procedure and consider only the cluster-level randomization of treatment. This increases the $n$ size by analyzing all clusters for which both participants returned to complete the experiment. However, the number of full clusters across conditions now varies with 84 clusters in the no contact condition, 77 clusters in the non-political contact condition, and 77 clusters in the political contact condition.

Table 6: Significance tests of average treatment effect estimates, outcome change in outparty affect, with full clusters

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-political contact vs. no contact</td>
<td>8.41</td>
<td>≈0</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>6.43</td>
<td>≈0</td>
</tr>
<tr>
<td>Non-political vs. political contact</td>
<td>1.98</td>
<td>.28</td>
</tr>
</tbody>
</table>

Note: Results considering all full clusters. Difference-in-differences estimates of average treatment effect and randomization inference $p$-values under the sharp null hypothesis of no treatment effect. Dependent variable is individual-level change in outparty affect. Two-sided $p$-values are reported. Both non-political contact and political contact improved outparty affect relative to no contact.

Table 7: Significance tests of average treatment effect estimates of outparty stereotypes, with full clusters

<table>
<thead>
<tr>
<th></th>
<th><em>Patriotic</em></th>
<th><em>Closeminded</em></th>
<th><em>Intelligent</em></th>
<th><em>Hypocritical</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>p-value</td>
<td>Estimate</td>
<td>p-value</td>
</tr>
<tr>
<td>Non-political contact vs. no contact</td>
<td>.29</td>
<td>.028</td>
<td>-.30</td>
<td>.008</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>.28</td>
<td>.027</td>
<td>-.28</td>
<td>.016</td>
</tr>
<tr>
<td>Non-political contact vs. political contact</td>
<td>.013</td>
<td>.96</td>
<td>-.02</td>
<td>.92</td>
</tr>
</tbody>
</table>

Note: Results considering all full clusters. Randomization inference hypothesis tests with the difference in means test statistic under the sharp null hypothesis of no treatment effect. Dependent variable is agreement, on a five point scale with higher values indicating more agreement, with how well the trait describes supporters of the outparty. Two-sided $p$-values are reported.
Appendix J  Robustness with other hypothesis test approaches

As a robustness check, I test my hypotheses by calculating $p$-values by using standard errors and $t-$values for matched-pair clustered designs (Blair 2019; Imai et al. 2009). Results are consistent with randomization inference results.

Table 8: Significance tests of average treatment effect estimates, outcome change in outparty affect

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. error</th>
<th>t-value</th>
<th>p-value</th>
<th>Conf. interval</th>
<th>Df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-political contact vs. no contact</td>
<td>9.17</td>
<td>1.80</td>
<td>5.08</td>
<td>$\approx 0$</td>
<td>[5.53, 12.81]</td>
<td>43</td>
</tr>
<tr>
<td>Political contact vs. no contact</td>
<td>7.47</td>
<td>2.40</td>
<td>3.11</td>
<td>.003</td>
<td>[2.62, 12.31]</td>
<td>43</td>
</tr>
<tr>
<td>Non-political vs. political contact</td>
<td>1.70</td>
<td>2.85</td>
<td>.60</td>
<td>.55</td>
<td>[-4.04, 7.45]</td>
<td>43</td>
</tr>
</tbody>
</table>

Note: Results considering all full blocks. Difference-in-differences estimates of average treatment effect and randomization inference $p$-values under the sharp null hypothesis of no treatment effect. Dependent variable is individual-level change in outparty affect. Two-sided $p$-values are reported. Both non-political contact and political contact improved outparty affect relative to no contact.