# Politics that Divide and Ties that Bind: Family, Friends, and Neighbors in a Polarized Era

Adam Smiley

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Cheryl Kaiser, Chair

Sapna Cheryan

Yuichi Shoda

Program Authorized to Offer Degree:

Psychology

University of Washington

## Abstract

Politics that Divide and Ties that Bind: Family, Friends, and Neighbors in a Polarized Era

Adam Smiley

Chair of the Supervisory Committee:

Cheryl Kaiser

Department of Psychology

Affective polarization has increased in the United States over the last several decades, leading to negative consequences for relationships, communities, and democracy. Why is affective polarization rampant, and what are the implications? In the first article of this dissertation, I examine a potential cause of affective polarization in the US: geographic sorting. We demonstrate that when people live in areas where they have little exposure to the out-party, they report generally higher levels of affective polarization. We follow this up by experimentally manipulating cross-party contact for students from politically homogenous universities and find that cross-party contact can reduce affective polarization for those in political bubbles. We argue that this cross-party contact is important for reducing affective polarization, and that increased political homogeneity within communities may have led to rising affective polarization in the United States.

In the second article of my dissertation, we examine whether perceived political difference is associated with worse relationship quality between college students and their parents. We find that when controlling for covariates such as non-political difference, parents' perceptions of political difference from their child is not associated with worse relationship quality. However, greater political difference was associated with worse relationship quality for students in some cases, but the effects were not large enough to be considered significant based on our hypotheses using minimum effect testing. In the third article of my dissertation, we propose a framework for statistical inference using minimum effect testing and equivalence testing. We use these methods in article two. This proposed framework allows researchers to test more specific hypotheses with a simple and versatile method that can be used for research questions.

Collectively, my dissertation furthers our knowledge of the causes and implications of affective polarization, as well as providing a statistical framework that can improve statistical inferences and research practices in quantitative research.

**Contents**

Introduction……..**6**

Article 1: Across the Aisle and Over the Fence: Geographic Sorting, Cross-Party Contact, and Affective Polarization……**12**

Article 2: Family Ties: Perceived College Student-Parent Political Differences and Relationship Quality……**67**

Article 3: Large enough to matter, too small to be of consequence, or insufficient evidence? A unified framework for a more meaningful statistical inference……...**107**

Conclusions and Future Directions……...…**134**

References……**138**

Acknowledgements……..**150**

Appendices

Article 1 Appendices……**152**

Article 2 Appendices……**160**

Article 3 Appendices……**169**

## Introduction

Affective polarization, defined as hostility and social distance between people from different political parties, has been steadily increasing in the United States since the 1980's (Mason, 2018; Iyengar et al., 2019). The United States is exceptional in this regard, with the rise in affective polarization outpacing many other developed countries during this period (Boxell et al., 2020). Today, most American partisans feel cold towards the *out-party* (the major party they do not identify with; Iyengar et al., 2019) and many have very few friends from across the partisan divide, if they have any at all (Cox, 2021). McCoy and Press (2022) argue that "pernicious polarization" in the United States may also be a risk to American democracy. Widespread conspiracy theories about the results of the 2020 Presidential Election (e.g., Pennycook & Rand, 2021) and the ensuing riots at the US Capitol on January 6th, 2021, suggest that perhaps polarization has already taken its toll on democracy in the US. To make matters worse, millions of Americans endorse political violence including assassinating leaders from opposing parties (Kalmoe & Mason, 2022).

How did we get here, and is there any reason for hope? In this dissertation, I identify a potential cause of affective polarization, geographic sorting, and suggest that this factor may point us to cross-party contact as a potential solution. Additionally, I examine the implications of this polarized era in the United States for families – and find that the parent-child relationship may be strong enough to largely withstand the forces of political prejudice.

### Part I: Geographic Sorting and Cross-Party Contact

Affective polarization is likely the result of many factors. Those who focus on institutional dynamics in the United States point to an entrenched two-party system without the emergence of a viable third party, as well as strong powers for the minority party such as the

filibuster (McCoy & Press, 2022). While these factors are certainly relevant, a social psychological lens is necessary to see the full picture. Political partisanship has become an important social identify for many Americans (Huddy & Bankert, 2017; Greene, 1999; West & Iyengar, 2020; Tajfel & Turner, 1979). As a result, research on social identities and intergroup relations may provide important insights for why affective polarization has become a powerful force in the United States.

Contact between members of different social groups is important for shaping attitudes (e.g., Allport, 1954; Pettigrew & Tropp, 2006, 2008). Intergroup contact between rival groups generally leads to more favorable evaluations of the outgroup (Pettigrew & Tropp, 2006) by reducing anxiety, increasing empathy, and increasing knowledge about the outgroup (Pettigrew & Tropp, 2008). Recent research has found this is often the case with political identities as well (Wojcieszak & Warner, 2020; Amsalem et al., 2021; Levendusky & Stecula, 2021; Santoro & Broockman, 2021). These findings suggest that a *lack* of cross-party contact may be related to higher levels of affective polarization. Might lower levels of cross-party contact be partially responsible for surging affective polarization in the United States?

Communities within the United States have become increasingly politically homogenous since the 1970s (Bishop & Cushing, 2008; Aisch et al., 2016; Bishop, 2020). Today, many cities, neighborhoods, and even individual streets are segregated by political party, with many Americans living in areas almost completely isolated from out-party members (Brown & Enos, 2021). In the first article of my dissertation, we (myself and Cheryl Kaiser) conduct two studies to examine if people who live in areas where they have limited exposure to the out-party are particularly likely to be affectively polarized. To test this, we determine respondents' level of cross-party exposure based on voting data for their communities of residence (counties and

precincts) and examine how cross-party exposure predicts affective polarization. We find that more exposure to the out-party is linked to lower levels of affective polarization (with one exception). These effects persist even when controlling for important covariates such as the strength of individuals' partisan identity and the policy views that they hold.

In study three, we built upon the first two studies by using an experiment. We created Democrat-Republican dyads over video chats with Democrats recruited from the liberal University of Washington and Republicans from Biola University and Liberty University, which are more conservative. We experimentally manipulated whether the dyads engaged in a contact activity (Fast Friends Protocol; Aron et al., 1997) or a non-contact activity (control), and we measured their levels of affective polarization before and after the intervention. We found that participants in the contact condition had larger reductions of affective polarization than those in the control condition, suggesting that intergroup contact caused a reduction in affective polarization. This is not the first experiment to examine the relationship between contact and affective polarization (e.g., Levendusky & Stecula, 2021; Santoro & Broockman, 2021), but this study provides several advantages for our research question. By specifically recruiting participants from politically homogenous universities, we were able to determine that contact can be effective for those who live in political bubbles. This is an important finding, as many people do live in politically homogenous communities (Brown & Enos, 2021) and our research suggests that those who are most isolated from the out-party are likely to have higher levels of affective polarization to begin with.

This research adds to our collective knowledge of why affective polarization occurs, and hints to contact as a means by which it might be reduced. One domain in which contact with people that hold different political views may occur is within families. Is this a great opportunity

for intergroup contact to occur? Or are politics a risk to family harmony? In Part II, I examine whether political differences can harm one of the strongest relational bonds: kinship.

**Part II: Perceived Political Difference in Parent-Child Relationships**

Even as offspring enter into adolescence and adulthood, the parent-child relationship is important both practically and psychologically. Young adults turn to their parents for help when making big decisions or handling emergencies (Pizzolato & Hicklen, 2011; Cooney & Uhlenberg, 1992), and many still receive financial support (e.g., Martin, 2019; Dickler, 2018). Parents may expect to receive practical support from their as they age, with adult children often fulfilling caretaking needs (Schulz et al., 2016; Ambert, 1997). For young adults, having positive relationships with their parents this can contribute to higher self-esteem (Noller & Callan, 1991), while both parents and their children experience less psychological distress when they have a close relationship (Umberson, 1992; Barnett et al., 1992). In sum, parents and their children have strong incentives to maintain a harmonious relationship when possible, and the parent-child bond is particularly strong relative to other forms of relationships (Amber, 1997; Laursen, 1993). In the second article of my dissertation, we (myself and Cheryl Kaiser) examine the link between perceived political differences and relationship quality for college students and their parents. Because of the solidity of the parent-child bond, this provides a stringent test case for the power of affective polarization.

Across three studies, we surveyed college students and their parents about their relationship with one another and how similar or different they perceived their political views to be. We measured three elements of relationship quality: interpersonal closeness, support, and negative interactions. In all three studies, we found that all three of these measures (zero-order) correlated with perceived political difference, such that greater difference was associated with

worse relationship quality. However, when including covariates such as non-political difference, the relationship between perceived political difference and relationship quality becomes less pronounced. We find no evidence that parents' perceived political difference is associated with any measure of relationship quality. For student respondents, we find a significantly non-zero effect of perceived political difference in some cases. Yet, these effects are quite small, and they were not sufficiently large to support our more conservative hypothesis based upon the principles of minimum effect testing.

In the third article of my dissertation, we (myself, Jessica Glazier, and Yuichi Shoda) suggest that minimum effect testing (Murphy & Myors, 1999; Serlin & Lapsley, 1985) and equivalence testing (e.g., Lakens et al., 2018; Tryon, 2001) be use in a unified framework for statistical inference. We show that these two tests are complimentary to one another, and we also provide a means by which minimum effect testing and equivalence testing can be performed simply using confidence intervals. Using this framework has several advantages over null hypothesis significance testing as it is traditionally practiced, which tests only if an effect can be considered significantly different from zero. Our proposed framework prompts researchers to consider what effect sizes would be meaningful a priori, and then uses those values to create a more specific null hypothesis based on a *null region*, rather than a single value. This framework is particularly useful for registered reports, where articles are considered for publication before data is collected. It is also useful for correlational research like ours.

In our investigation in article two, our aim was to pinpoint the association between perceived political difference and relationship quality. We included several covariates, including non-political difference, in order to try to control for the myriad other factors that influence the quality of parent-child relationships. However, even with these covariates, there are likely other

factors that covary with perceived political difference which may also influence relationship quality. By using minimum effect testing based on the crud factor (Orben & Lakens, 2019), we made our hypothesis testing more conservative to account for this noise associated with our primary predictor variable which may amplify the effect. Reporting our results in terms of minimum effect testing and null hypothesis testing provided nuance to our conclusions when we observed non-zero relationships (as in Study 1): we determined that there is an association between perceived political difference and relationship quality, but the effect was not large enough to conclude that it was large enough to matter for our research question.

**ARTICLE 1**

# Across the Aisle and Over the Fence: Geographic Sorting, Cross-Party Contact, and Affective Polarization

Adam H. Smiley and Cheryl R. Kaiser

University of Washington

## Author Note

This project is supported in part by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP).

Correspondence concerning this article should be addressed to Adam H. Smiley, University of Washington, 119A Guthrie Hall Box 351525, Seattle, WA 98195-1525.

Email: ahsmiley@uw.edu

Abstract

Across three studies, we examine the role of geographic sorting on affective polarization. In Study 1, we analyze data from the 2016 American National Election Study (ANES) and find that Americans who live in counties with greater proportions of out-partisans tend to have lower levels of *party*-directed affective polarization, even after controlling for party identity strength and policy views. In Study 2, we conducted a pre-registered study with original data and find that partisans who have more cross-party exposure in their counties and precincts report lower levels of social distance from out-party members, but not significantly lower *voter*-directed affective polarization. Finally, a controlled experiment with Democrat-Republican dyads from politically homogenous universities demonstrated that cross-party contact reduces affective polarization and social distance for individuals with little cross-party exposure.

I don't see red states and blue states, but only United States.

−President Elect Joe Biden, November 7th, 2020

From states, to counties, to neighborhoods, it is hard not to see the United States as clusters of red and blue. Voter registration data from 180 million Americans revealed that more than half of all American partisans live in a neighborhood[1] with less than 33% members from the out-party[2] (Brown & Enos, 2021). Even more strikingly, one in ten Democrats lives in a neighborhood composed of less than 5% Republicans, and one in ten Republicans lives in a neighborhood with less than 10% Democrats. In the 2020 US presidential election, 79% of counties were won by 20 percentage points or more (Bishop, 2020), reflecting a huge increase in geographic sorting since 1976, in which only 36% of counties were won by the same margin in a closely contested presidential election (Bishop & Cushing, 2008). We examine how geographic sorting of political partisans in the United States relates to *affective polarization*, defined as animosity and social distance between people of different political parties (Iyengar et al., 2019; Mason, 2018). In particular, we investigate whether partisans who are exposed to fewer out-party supporters in their communities are more affectively polarized. Furthermore, an important implication of geographic sorting is that those living in politically homogenous areas aligned with their views have less intergroup *contact* with out-partisans. Using an experiment, we examine whether cross-party contact can reduce affective polarization and facilitate cooperation.

As geographic sorting of American partisans has increased over the last 50 years, affective polarization has also increased (Iyengar et al., 2019; Mason, 2018). The US has seen a larger uptick in affective polarization than many other developed countries during this timeframe

[1] Brown and Enos (2021) used a weighted analysis of the 1000 closest neighbors to each registered voter, with those closer in distance receiving more weight.

[2] Throughout this paper, we refer to members of an individual's political party as "in-party" and a member of a different political party as "out-party." For example, for a Democrat, the Democratic Party would be considered their in-party while the Republican Party (or GOP) would be considered their out-party.

(Boxell et al., 2020). The rise in affective polarization is evident in explicit measures, such as feeling thermometers (Iyengar et al., 2019; Mason, 2018) and behavioral reports (e.g., friendships and romantic relationships across political lines, Huber & Malhotra, 2017) as well as on implicit measures (Iyengar and Westwood, 2015). Beyond simply disliking and avoiding partisans across the aisle, the share of Americans endorsing violence against people from the opposing political party has also increased (Kalmoe & Mason, 2019, 2022). Frequent federal government shutdowns and the January 6th Capitol riot suggest that affective polarization may be a threat to effective governance and democracy. While some political animosity can be attributable to policy disputes (i.e., *ideological* polarization), even partisans with moderate political views can harbor extreme levels of political hostility (Mason, 2018).

Individuals without strong policy positions can hold strong prejudices towards out-partisans in large part because political parties have become important social identities (Huddy & Bankert, 2017; Greene, 1999; West & Iyengar, 2020; Tajfel & Turner, 1979). One way that individuals attempt to maintain self-esteem is through comparing their social groups to others (Lemyre & Smith, 1985). Even when group memberships are arbitrary, ingroup favoritism and outgroup bias can arise (e.g., Tajfel, 1970; Diehl, 1990). Yet, political parties are far from minimal or arbitrary groups. In the context of the United States' two-party system, Democrats and Republicans are in zero-sum competition for electoral success, with a victory for the out-party generally at the expense of the in-party (Wojcieszak & Warner, 2020). Additionally, political partisanship overlaps considerably with other meaningful social identities such as race, ethnicity, and religion (Huddy & Bankert, 2017; Mason, 2018). Convergence of multiple group memberships is related to increased intolerance of outgroup members (Roccas & Brewer, 2002),

and this "social sorting" within political parties has also contributed to a rise in affective polarization (Mason, 2018).

With political partisanship functioning as a salient social identity for many Americans, a social identity lens may be helpful for understanding why affective polarization has increased and how it might be reduced. An extensive literature suggests that intergroup contact can serve to reduce prejudice between people from different groups, and recent research suggests that contact between partisans is no exception. The effect of contact on intergroup attitudes is particularly pertinent to our research question, given that increasing political homogeneity within communities inevitably results in less contact between people from different political parties.

Allport (1954) suggested that contact could reduce prejudice between groups under four conditions: (1) equal status between groups in the situation, (2) common goals, (3) institutional support such as laws and norms, and (4) intergroup cooperation. A meta-analysis of over 500 studies found that in most cases intergroup contact reduces prejudice between groups (Pettigrew & Tropp, 2006). The effect of contact on prejudice is mediated by reduced anxiety, increased empathy, and, to a lesser degree, increased knowledge of the outgroup (Pettigrew & Tropp, 2008). However, Pettigrew and Tropp (2006) concluded that Allport's four conditions are not *necessary* for intergroup contact to reduce prejudice, but they do increase the likelihood that contact will have positive results (see also Pettigrew et al., 2011). This is important in the domain of affective polarization, as political partisans often lack common goals and are frequently competing (rather than cooperating) with one another (Mason, 2018).

Contact between political partisans has also been linked to reduced affective polarization. Individuals with more out-party friends reported lower levels of affective polarization (Wojcieszak & Warner, 2020). Although this relationship is likely bidirectional, in this study

having friends-of-friends from the out-party (i.e., "extended contact") also predicted lower affective polarization. Canadians with more frequent cross-party discussions (either in person or online) *about politics* reported lower levels of affective polarization (Amsalem et al., 2021). In contrast, experiments conducted with US samples have found that cross-party face-to-face discussions can reduce affective polarization, though only when affective when partisans were encouraged to talk about common ground between the parties (Levendusky & Stecula, 2021) or when discussing their perfect days (Santoro & Broockman, 2021), but not when discussing polarization (Santoro & Broockman, 2021). Taken together, the growing body of research on contact and affective polarization suggests that cross-party contact has the potential to reduce partisan prejudice, but there is likely nuance to the types of contact that have this effect and perhaps even the types of people for whom contact reduces polarization.

If cross-party contact is important for reducing affective polarization, the increased political homogeneity of American communities may have played a role in the rise in affective polarization over the last 50 years. Bishop and Cushing (2008) detail the rise of geographic sorting in recent decades and suggest that this has caused an increase in affective polarization. However, they do not empirically test how living in a politically homogenous place relates to *individuals'* perceptions of the out-party. Rather, they focus broadly on nationwide trends. In contrast to much (though not all) of political science research, psychological research focuses on individual level processes. There are several advantages to examining the relationship between cross-party exposure and affective polarization from an individual level. If individuals living in localities with limited cross-party exposure really are more affectively polarized, this provides important evidence that the political makeup of a person's community is related to how people perceive political partisans. This has implications for understanding what drives affective

polarization. Additionally, examining the causal effects of cross-party contact for individuals with limited cross-party exposure may elucidate a potential means by which affective polarization can be reduced.

Across three studies, we examine the role that cross-party contact plays in shaping attitudes towards partisans. We use a mixed-methods approach in order to examine the extent to which cross-party exposure predicts affective polarization, and then to examine whether cross-party contact can reduce affective polarization, particularly for those who have minimal prior cross-party exposure. In Study 1, we used a large nationwide dataset to examine how exposure to out-partisans at the county level is associated with affective polarization. In Study 2, we surveyed participants to examine the relationship between cross-party exposure (at the county and precinct levels) two different measures of affective polarization. In Study 3, we use an experiment to test if cross-party contact can reduce affective polarization for students at politically homogenous universities.

## Study 1

Study 1 uses data from the 2016 American National Election Study (ANES; American National Election Studies, 2017) to examine whether county-level exposure to members of the out-party predicts lower levels of affective polarization. Based on the extensive literature linking intergroup contact with reduced prejudice, we hypothesize that higher levels of cross-party exposure will predict lower levels of affective polarization.

Counties are relatively large, with average populations over 100,000. Within a single county, politically homogenous neighborhoods that are primarily Republican and others that are mostly Democratic are common (Kaplan et al., 2020). County is the most precise level of

analysis available for the ANES that can be mapped onto election results,[3] but given the amount of within-county variation on cross-party exposure, county-level exposure is an imprecise measure of cross-party exposure. Thus, the anticipated noise in the data makes this a particularly stringent test for our theory. As a result, we consider any significant non-zero relationship between cross-party exposure and affective polarization to be theoretically meaningful (Smiley et al., under review).

## Method

### *Participants*

We analyzed responses from the 2016 ANES (American National Election Studies, 2017) and associated location data from the restricted-access geocodes (American National Election Studies, 2021). There were 4270 respondents before exclusions. However, we only included responses of those who identified with the Democratic or Republican parties, or who identified as independent but leaned towards the Democratic or Republican parties (14% excluded). Additionally, we excluded participants who did not provide an answer for one or more of the measures that we included in our model (see Fig. 4; additional 31% excluded). In total, we included responses from 2355 participants (51.6% women, 48.4% men; $M_{\text{Age}}$ = 49.78, $SD_{\text{Age}}$ = 17.34). The sample was relatively even in the number of people that identify with or lean towards the two major political parties, with 52.0% Democratic party identifiers/leaners and 48.0% closer to the Republican party. The three most common racial/ethnic identities in the sample were White (75.6%), Black (8.9%), and Hispanic (8.5%). Participants came from 795 different counties in the United States.

### *Measures*

[3] Zip codes are collected, but zip code level voting results are not readily available.

The measures used in this analysis were based on a model used by Mason (2018) to examine the effects of social sorting on affective polarization. Because this analysis was not pre-registered, we linked it as closely as possible to this model previously used to predict affective polarization. With two exceptions, our model uses the same variables as Mason.[4] First, rather than examining social sorting, our main variable of interest is cross-party exposure at the county level. And second, Mason uses issue extremity and issue constraint (the degree to which views on different topics are similar in how liberal/conservative they are) to collectively measure how aligned individuals' political views are with their political parties, which we do with just one variable (issue polarization). Issue polarization captures the extent to which individuals hold more extreme political views consistent (or inconsistent) with their political parties.

***Affective Polarization (Party-Directed)[5].*** Affective polarization was assessed with feeling thermometers, a common way of measuring affective polarization (Druckman & Levendusky, 2019). Participants are asked how warmly they feel towards each major political party on a scale of 0 (very cold) to 100 (very warm). To calculate an individual's level of affective polarization, participant's feeling thermometer rating of their in-party are subtracted from the feeling thermometer rating of their out-party. For example, if a Democrat's feeling thermometer rating for Democrats is 70 and rating for Republicans is 20, their level of affective polarization would be 50.

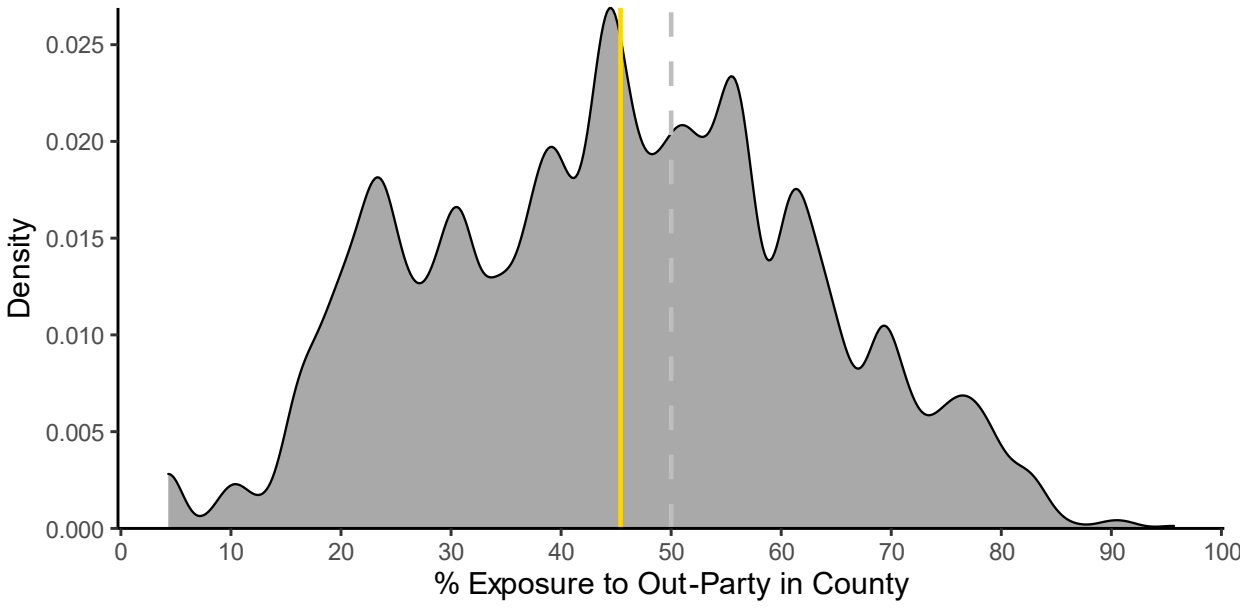
***Cross-Party Exposure.*** We determined a participant's level of cross-party exposure based on the proportion of voters in the participant's county who voted for the 2016 presidential candidate from the major party (Democrat/Republican) that the participant *does not* identify with. County level voting data was obtained from the MIT Election Data and Science Lab

[4] In Figure 5.2 and its originating regression in Table A.5.
[5] In studies 2 and 3 we also examine "voter-directed" affective polarization, that asks about warmth towards party voters rather than parties as a whole.

(2018). Third-party and write-in votes were excluded in this analysis, so that the "total" votes for a given county were the sum of votes for the Democratic party candidate (Hillary Clinton) and the Republican party candidate (Donald Trump). To demonstrate, if a Republican (or a Republican leaner) lived in a county where Hillary Clinton received 70% of the votes for major party candidates, then that individual's level of cross-party exposure would be 70%. However, if the participant was a Democrat (or Democrat leaner) in the same county, their level of cross-party exposure would be 30%. See Fig. 1 for a breakdown of participants' levels of cross-party exposure. This variable, like all other predictor variables in this analysis, was rescaled to a 0-1 scale so that 0 represents the smallest value (least exposed participant) and 1 represents the highest value (most exposed participant).

**Figure 1**

*Density Plot of Participants' Cross-Party Exposure in their County of Residence.*




*Note*. Unscaled values. The dashed grey line is at 50%, and the yellow line represents the median cross-party exposure.

***Issue Polarization.*** This measure reflects how closely respondents' policy views align with the more extreme views of their political party.[6] Issue polarization measured the extent that participants' views on five policies were consistent with their political party. The five topics, which were selected to get a wide range of different types of policies, were: (1) government provided services, (2) government spending on healthcare, (3) citizenship/deportation for illegal immigrants, (4) abortion, and (5) government help to Black-Americans. They were coded such that if a participant selected the most extreme position that was in line with their political party it was scored as a 1. If the participant indicated support for the most extreme position in line with the opposing political party, this would be scored as a 0. Policy positions not at the extremity were spaced evenly between 0 and 1 (there were four to seven options per question). To illustrate, if a Democrat indicated that abortion should be legal under all circumstances, they would score 1 for that question, as the Democratic party is generally pro-choice. A Republican would score 0 if they indicated they supported abortion under all circumstances, as the Republican party is generally pro-life. Responses on all five questions were averaged, so that those scoring high on the 0-1 scale held policy views generally in line with their political party, while those who scored low held views generally aligned with the other major political party.

***Party ID Strength.*** This variable represents the extent to which participants identify with their political party. A value of 0 represents an independent that leans towards a party, .5 represents a moderate partisan (Democrat/Republican), and 1 represents a strong partisan.

***Education.*** Level of education, with higher values representing more formal education. The five levels are (0) did not complete high school or GED, (.25) high school graduate or GED,

[6] This is similar to the concept of "operational ideology" (Ellis & Stimson, 2012).

(.5) associates degree or some college, (.75) bachelor's degree, and (1) graduate or professional degree.

***Political knowledge.***[7] This variable represents the proportion of four political knowledge questions that were answered correctly, with a maximum score of 1 (all answers correct) and a minimum of 0 (all answers incorrect). The four questions (and correct answers) were the length of a senate term (6 years), party with the majority in the House of Representatives in 2016 (Republicans), party with the majority in the Senate in 2016 (Republicans), and which of the following receives the least funds from the federal government: foreign aid (correct answer), social security, defense spending, Medicare.

***White.*** Dummy coded variable with a score of 1 representing someone who identified as White (non-Hispanic) and 0 representing any other race/ethnicity.

***Hispanic.*** Dummy coded variable with a score of 1 representing someone who identified as Hispanic and a 0 representing those who do not identify as Hispanic.

***Black.*** Dummy coded variable with a score of 1 representing someone who identified as Black (non-Hispanic) and a 0 representing any other race/ethnicity.

***Male.*** Dummy coded variable for the participant's sex, male (1) or female (0). All participants in the sample identified as either male or female.

***Income.*** Household income designated as within one of 28 ranges. The lowest range, $5000 or less, was scored as 0, and the highest range, $250,000 or more scored, as 1.

***Age***. Age in years. Respondents were at least 18 years of age. Respondents over 90 years of age were listed as "90 or above". Values were recoded to a 0-1 scale such that 0 represented 18-year-olds and 1 represented those 90 years of age or older.

[7] Mason (2018) instead uses the term "sophistication" to refer to political knowledge.

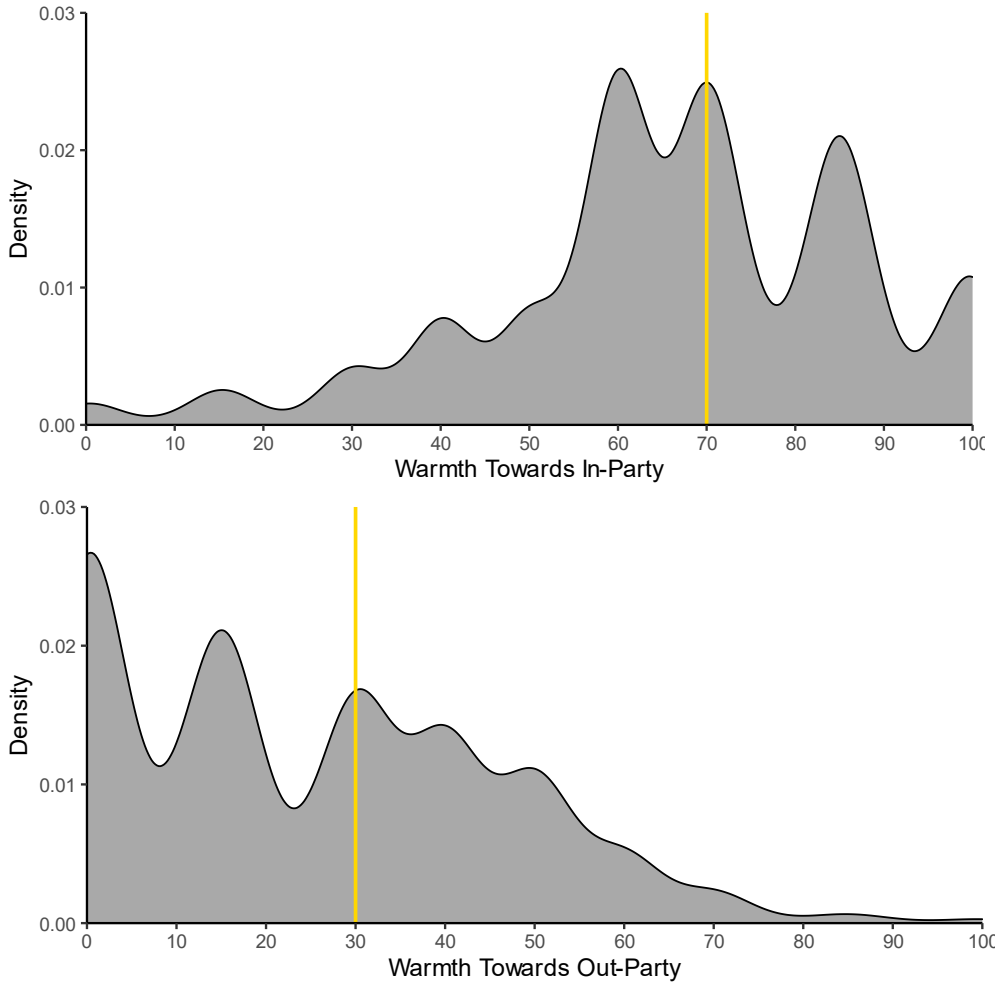
***Church attendance.*** This variable represents the frequency with which an individual attends church or religious ceremonies. Higher scores represent more frequent church attendance. The five levels are (0) never, (.25) a few times a year, (.5) once or twice per month, (.75) almost every week, and (1) every week.

## Results

The mean level of affective polarization for participants was 41.79 (*SD* = 31.32), reflecting warmer feelings towards the in-party (*M* = 66.48, *SD* = 21.03) than the out-party (*M* = 24.69, *SD* = 21.07). As demonstrated by the density plots in Fig. 2, a large number of participants rated the out-party as 0 out of 100 on the feeling thermometer.

**Figure 2**

*Density plots of in-party and out-party feeling thermometer ratings in Study 1.*

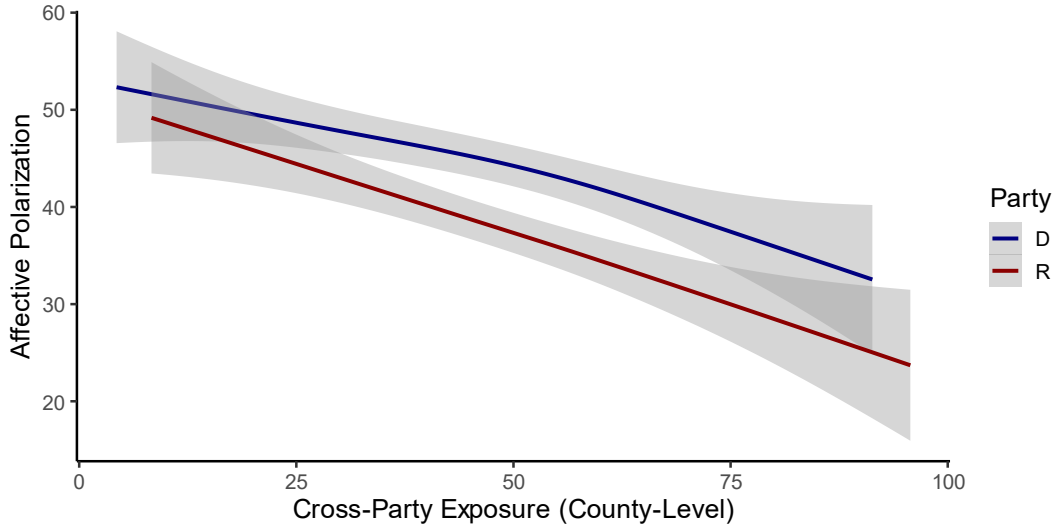



*Note*. The yellow line represents the median.

We first calculated the zero-order correlation between cross-party exposure at the county level and party-directed affective polarization. There was a significant negative relationship between the two variables ($r = -.15$, $t(2353) = -7.22$, $p < .001$, 95% CI = [−.19, −.11]), suggesting that those who lived in counties with more members of the out-party had lower levels of affective polarization. As shown in Fig. 3, this was observed for both Democrats and Republicans.

**Figure 3**

*Relationship Between County-Level Cross-Party Exposure and Party-Directed Affective Polarization in Study 1.*

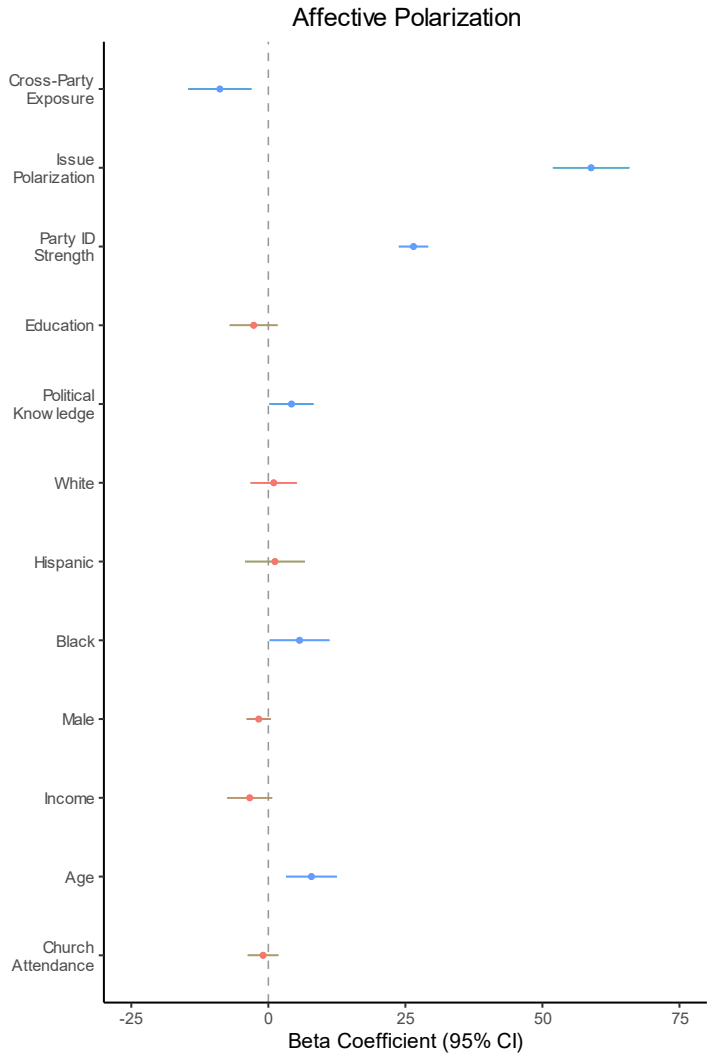



*Note.* Red and blue lines represent general additive models (GAM) for Republican and Democratic respondents respectively, with no covariates. The shaded area represents 95% CIs.

To control for other potentially important variables, we next conducted a multiple linear regression predicting an individual's level of party-directed affective polarization from county-level cross-party exposure and a number of covariates (see Fig. 4 for full model). In line with our hypothesis, we found a significant negative relationship between cross-party exposure and affective polarization, $\beta = -8.85$, $p = .003$, 95% CI = [−14.61, −3.08]. In other words, we would expect that respondents with very little cross-party exposure to be about nine points more affectively polarized than respondents with very high levels of cross-party exposure, all else being equal. This pattern is evident even when controlling for strong predictors that account for a huge proportion of the variance, notably issue polarization ($\beta = 58.88$, $p < .001$, 95% CI = [51.91, 65.86]) and partisan ID strength ($\beta = 26.48$, $p < .001$, 95% CI = [23.80, 29.16]).

**Figure 4**

*Coefficients for a Multiple Linear Regression Predicting Party-Directed Affective Polarization in Study 1.*

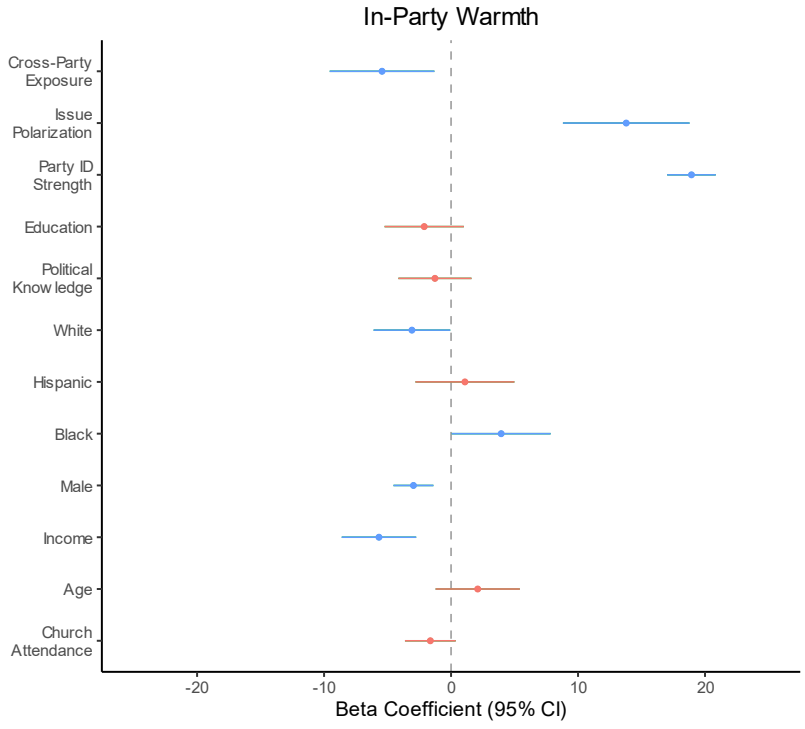



*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant. Affective polarization = in-party feeling thermometer rating – out-party feeling thermometer rating.

This measure of party-directed affective polarization is based on the difference between in-party evaluations and out-party evaluations. As a result, we can also look at how cross-party exposure relates to these individual components. To investigate these effects, we conducted multiple linear regression models predicting in-party evaluations and out-party evaluations, using the same predictors from the previous model. We find that more cross-party exposure predicts significantly *lower* warmth towards the in-party, $\beta = -5.43$, $p = .01$, 95% CI = [−9.56, −1.31] (see Fig. 5). However, we find that the relationship between cross-party exposure and out-party warmth is not statistically significant, $\beta = 3.41$, $p = .11$, 95% CI = [−0.71, 7.54] (see Fig. 6).

**Figure 5**

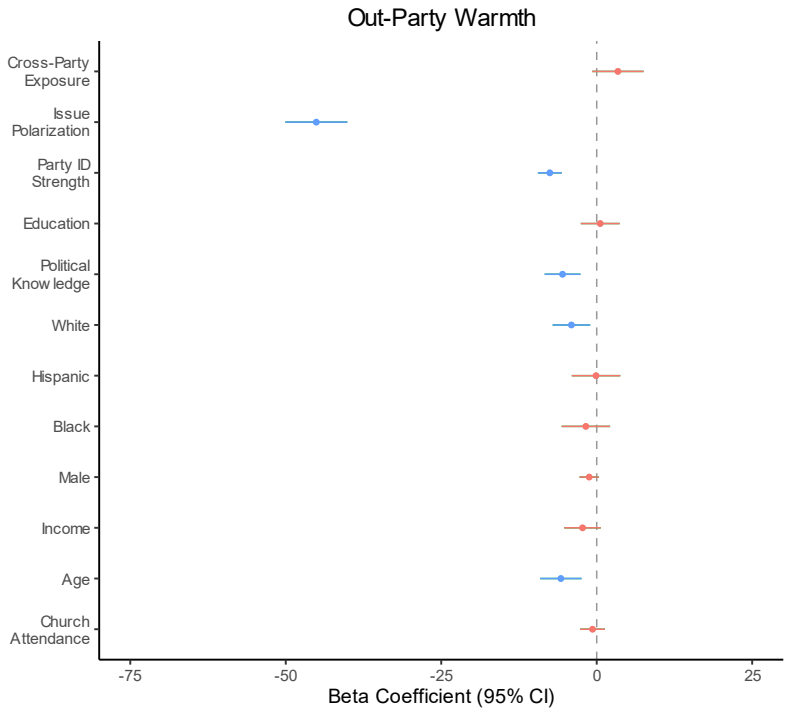
*Coefficients for a Multiple Linear Regression Predicting In-Party Warmth in Study 1.*



*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

**Figure 6**

*Coefficients for a Multiple Linear Regression Predicting Out-Party Warmth in Study 1.*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

**Discussion**

Consistent with our hypothesis, people with more exposure to out-partisans in their county had lower levels of party-directed affective polarization. Importantly, this effect holds even when controlling for important individual-level variables such as issue polarization and partisan identity strength. Using county level data is relatively imprecise, and as a result presents a stringent test of this theory. However, conducting this analysis based on cross-party exposure with a smaller unit of measurement than the county would be beneficial as well, because partisans can become clustered in small areas such as precincts or even single streets (Brown & Enos, 2021; Kaplan et al., 2020).

It is noteworthy that the relationship between cross-party exposure and affective polarization was the result of an increase in in-party warmth to a greater degree than a decrease in out-party warmth. Prejudice can manifest both as in-group favoritism and out-group hate, and the two can sometimes be independent of each other (Brewer, 1999). However, discrimination is often the result of ingroup favoritism rather than blatant outgroup hate (Greenwald & Pettigrew, 2014), and the data in Study 1 suggest that intergroup attitudes are improved by reducing ingroup favoritism.

In Study 2, we build upon the findings from Study 1 and include the precinct as an additional level of analysis for examining the relationship between cross-party exposure and affective polarization. We also add a measure of *social distance*, another component of affective polarization, which indicates how comfortable (or uncomfortable) individuals are having interpersonal relationships with out-partisans and may be particularly important for understanding behavioral outcomes such as discrimination (Druckman & Levendusky, 2019) and a willingness to interact with outpartisans.

## Study 2

In Study 2, we conducted a pre-registered study (Pre-Reg: https://osf.io/r52jt) using participants recruited through Prolific. We hypothesized that there would be a significant relationship between county-level cross-party exposure and voter-directed affective polarization as measured by feeling thermometers. In Study 2, we asked participants how warm (or cold) they felt towards the *voters* of the two major political parties, rather than the parties themselves. politicians, pundits, etc.) more so than non-elite partisans (Druckman & Levendusky, 2019). This subtle change in our measure from Study 1 allows us to examine the factors that influence how respondents feel towards these non-elite partisans they are likely to encounter in their daily lives. This likely has more of an impact on actual cross-party behaviors (excluding voting, campaign contributions, etc.), as most people rarely interact with political elites.

In Study 2 we also examined the relationship between *precinct*-level cross-party exposure and affective polarization. Partisans can cluster in small areas within counties (Brown & Enos, 2021; Kaplan et al., 2020), and as a result, someone who lives in a county where they are surrounded by mostly out-partisans may actually live in a precinct[8] that is comprised of mostly in-party members. We hypothesized that precinct-level cross-party exposure would also predict higher levels of affective polarization as measured by feeling thermometers.

We also extended our examination from Study 1 to include a new measure of affective polarization that focuses on social distance. This addition allows us to get a sense of how people feel about interacting with the out-party. This is important, as even those in very politically homogenous communities will come into contact with out-partisans at some point and will need to make decisions about whether to engage socially with them or not. Additionally, social distance measures capture something slightly different from feeling thermometer evaluations, as

[8] Precincts are the smallest unit that electoral districts can be divided into. On average, there are just over 1000 registered voters per precinct in the US.

it is possible to feel coldly towards a group of people but still feel comfortable having interpersonal relationships with them. We hypothesized that higher levels of cross-party exposure at both the county level and the precinct level would predict lower levels of social distance.

## Method

### *Participants*

We recruited 401 participants through Prolific in May of 2022 who initially identified as Democrats or Republicans. We excluded seven states from recruitment (AK, AL, ID, KY, LA, MO, and VA) because of a lack of precinct-level voting data in those states.[9] In line with our pre-registration, six participants were excluded for not identifying with the Democratic or Republican party, and 33 were excluded for not providing an answer for one or more of the measures that we included in our models (see Figs. 8-11). In total, we included responses from 362 participants (46.4% women, 51.6% men, 1.9% non-binary; $M_{\text{Age}}$ = 41.99, $SD_{\text{Age}}$ = 14.75), with an even number of Democrats and Republicans (181 each). The four most common racial/ethnic identities in the sample were White (77.9%), Asian (6.4%), Hispanic/Latino/a (5.2%), and Black (5.0%). Participants lived in 241 different counties in the United States.

### *Measures*

We used two main outcome variables in Study 2: voter-directed affective polarization (feeling thermometers) and social distance.

***Affective Polarization (Voter-Directed).*** Participants were asked how warm (or cold) they felt towards each major political party's *voters* on a scale of 0 (very cold) to 100 (very warm). This is slightly different from in Study 1, where feeling thermometers referred to the

[9] Results from absentee/mail-in ballots were reported at the county level in many of these states, rather than the precinct level, in 2020. The 2020 US Presidential election occurred during the COVID-19 pandemic, and as a result many voting and reporting procedures were altered.

parties as a whole, rather than voters. To calculate an individual's level of affective polarization, participant's feeling thermometer rating of their in-party voters are subtracted from the feeling thermometer rating of their out-party voters.

***Social Distance.*** Another way that affective polarization is often quantified is based on social distance measures (Druckman & Levendusky, 2019; see also Bogardus, 1933).[10] We asked participants how comfortable they would be having a member of the opposite political party as a (1) close friend, (2) romantic partner, (3) neighbor, and (4) a close coworker. They responded on a 1 (extremely uncomfortable) to 7 (very comfortable) scale. Responses about these four relationships were averaged into a combined social distance score for each participant. For analysis, we reverse coded this variable so that a higher score represented greater discomfort (i.e., social distance).

We had two primary predictor variables of interest, cross-party exposure at the county level (as in Study 1) and cross-party exposure at the precinct level.

***Cross-Party Exposure (County-Level).*** We determined a participant's level of cross-party exposure based on the proportion of voters in the participant's county who voted for the 2020 presidential candidate from the major party (Democrat/Republican) that the participant *does not* identify with. County level voting data was obtained from the MIT Election Data and Science Lab (2018). Third-party and write-in votes were excluded in this analysis, so that the "total" votes for a given county were the sum of votes for the Democratic party candidate (Joe Biden) and the Republican party candidate (Donald Trump). This variable was rescaled to a 0-1 scale so that 0 represents the smallest value (least exposed participant) and 1 represents the highest value (most exposed participant).

[10] Although social distance is a component of affective polarization more broadly, in our models we use the term "affective polarization" to refer to feeling thermometer ratings and "social distance" to refer to these social distance ratings.

***Cross-Party Exposure (Precinct-Level).*** Participants were instructed to follow a link to the *New York Times* "Extremely Detailed Map of the 2020 Election" (Park et al., 2021). They were instructed to enter their current address to find results for their precinct. They then reported the candidate who won in their precinct and his margin of victory. A cross-party exposure value was calculated from this information. When participants lived in precincts where their candidate *won*, their cross-party exposure score was the margin of victory multiplied by −1 to make the value negative. When participants lived in precincts where their candidate *lost*, their cross-party exposure score was simply the margin of victory in that precinct. As a result, more positive values reflect participants living in precincts with many voters for the out-party in the 2020 presidential election, while those with very negative values are living in areas with many voters from the in-party in 2020. Like other variables, this variable was scaled to a 0 (least exposed participant) to 1 (most exposed participant) scale before being used in regression models.

The predictor variables used in our pre-registered models were the same as those used in Study 1, with a few exceptions (see full models in Figs. 8-11):

***Democrat.*** We added a dummy coded variable for the participant's political party, with a 1 representing Democrats and 0 representing Republicans.

***Party ID Strength.*** In Study 2, we did not include partisan leaners so Party ID Strength is now dummy coded, with 0 representing a self-identified "not very strong" partisan and a 1 representing a "strong" partisan.

***Political knowledge.*** Questions are slightly different than in Study 1. The four questions (and correct answers) in Study 2 were the length of a senate term (6 years), party with the majority in the House of Representatives in 2022 (Democrats), the current Chief Justice of the

United States (John Roberts), and which of the following receives the least funds from the federal government: foreign aid (correct answer), social security, defense spending, Medicare.

***Man*.** In Study 2, we asked about participants gender identities rather than sex, so we used the term "man" rather than "male". Man is a dummy coded variable, with a score of 1 representing a man and a score of 0 representing a person with any other gender identity.

For exploratory purposes, We also asked about participants levels of *subjective cross-party contact* with the question "In your community, do you interact with more Democrats or more Republicans on a regular basis?" The scale ranged from 0 (Many more democrats) to 100 (many more Republicans). For analysis, we coded this variable so that a higher score represented more cross-party contact, and scaled the values between 0 and 1. We also asked several questions about participants relationships with in-party and out-party members. We asked participants how many "friends and good acquaintances" and family members they had from the two main political parties (none, a few, many, or very many. An attention check administered in the middle of the study instructed participants to select a specific answer choice if they were paying attention. See Supplemental Materials for complete question text.

**Results**

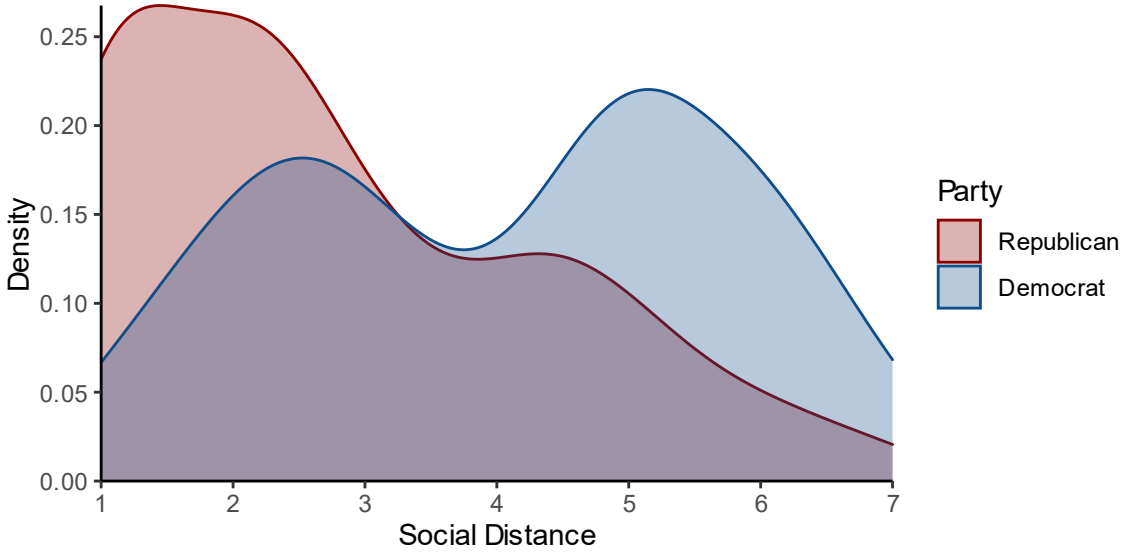
The mean level of voter-directed affective polarization for participants was 44.23 (*SD* = 31.36), reflecting warmer feelings towards in-party voters (*M* = 75.47, *SD* = 18.33) than out-party voters (*M* = 31.23, *SD* = 24.98). Voter-directed affective polarization was higher among Democrats (*M* = 49.69, *SD* = 30.54) than Republicans (*M* = 38.8, *SD* = 31.29), $t(359.8) = 3.35$, $p < .001$. The mean level of social distance was 3.43 (*SD* = 1.76), slightly below the midpoint on the 1 (very comfortable) to 7 (very uncomfortable) scale,[11] indicating that strong discomfort with having interpersonal relationships with out-partisans was not the norm. Social distance was also

[11] As described in the methods section, this variable was reverse coded.

higher for Democrats ($M$ = 4.08, $SD$ = 1.68) than Republicans ($M$ = 2.77, $SD$ = 1.58), $t$(358.68) = 7.61, $p$ < .001 (see Fig. 7). County- and precinct-level cross-party exposures were highly correlated ($r$ = .68).

**Figure 7**

*Density plots of social distance ratings by party in Study 2.*




*Note.* Democrat and Republican distributions are overlapping, not additive/stacked. Higher values denote greater social distance.

We calculated the zero-order correlation between cross-party exposure at the county level and voter-directed affective polarization as measured using feeling thermometers. We found no significant relationship between the two variables ($r$ = −.06, $p$ = .23). We also found that there was no significant zero-order correlation between precinct-level cross-party exposure and voter-directed affective polarization ($r$ = −.09, $p$ = .10).

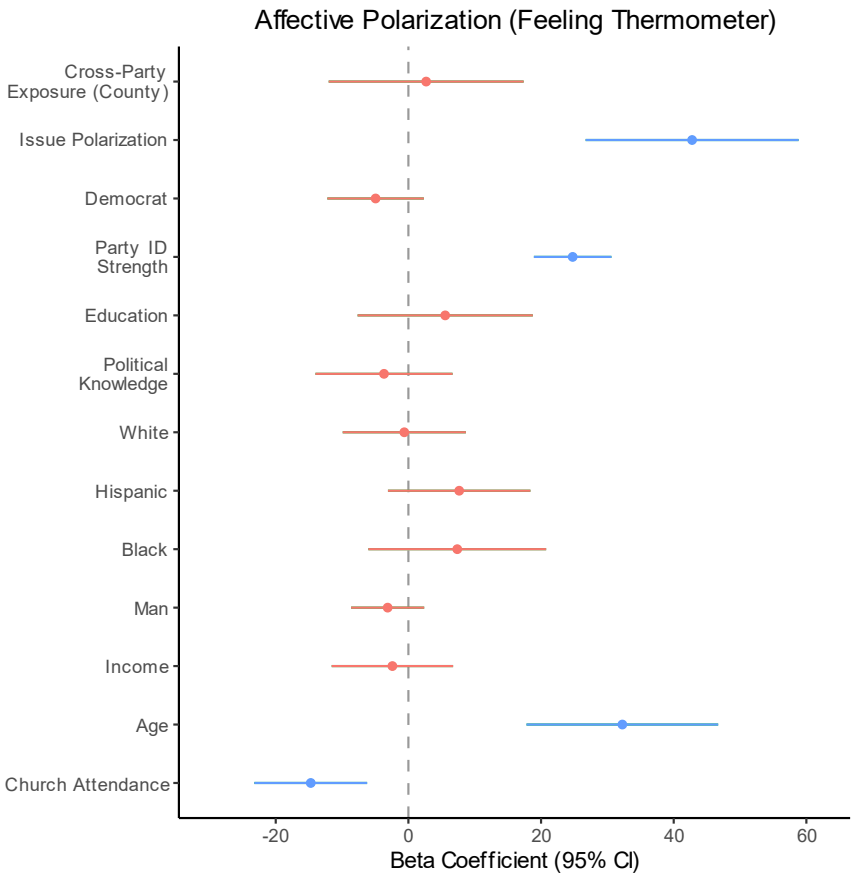
We also examined zero-order correlations between cross-party exposure and social distance. We found that there was a significant correlations using both county-level ($r = -.22$, $p < .001$) and precinct-level cross-party exposure ($r = -.21$, $p < .001$). This suggests that those who have more exposure to out-partisans in their county and in their precinct reported lower levels of social distance.

***Pre-Registered Analyses***

Based on our pre-registered hypotheses and analysis plan, we conducted four multiple linear regression models. First, we examined the relationship between affective polarization and county-level cross-party exposure as well as a number of covariates (see Fig. 8 for full model). Contrary to our hypothesis and the model in Study 1, we found no significant relationship between cross-party exposure and affective polarization, $\beta = 2.67$, $p = .72$, 95% CI = [−12.02, 17.37]. Next, we examined the relationship between *precinct*-level cross-party exposure and affective polarization. The same model using precinct-level cross-party exposure similarly found no significant relationship between cross-party exposure and affective polarization, $\beta = 0.40$, $p = .95$, 95% CI = [−12.34, 13.16] (see Fig. 9).

**Figure 8**

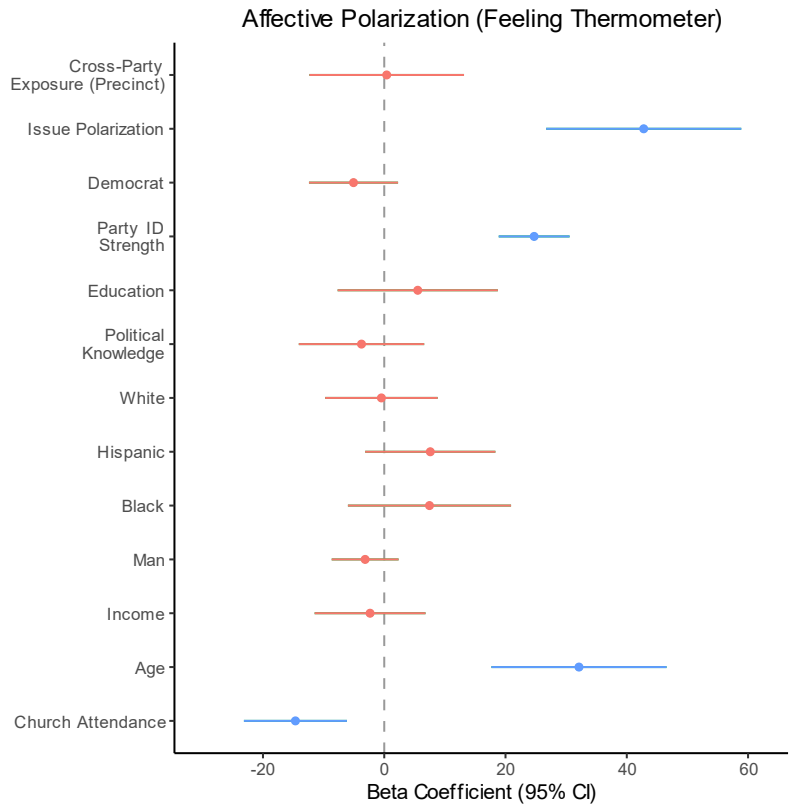
*Predicting affective polarization using county-level cross-party exposure in Study 2.*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

**Figure 9**

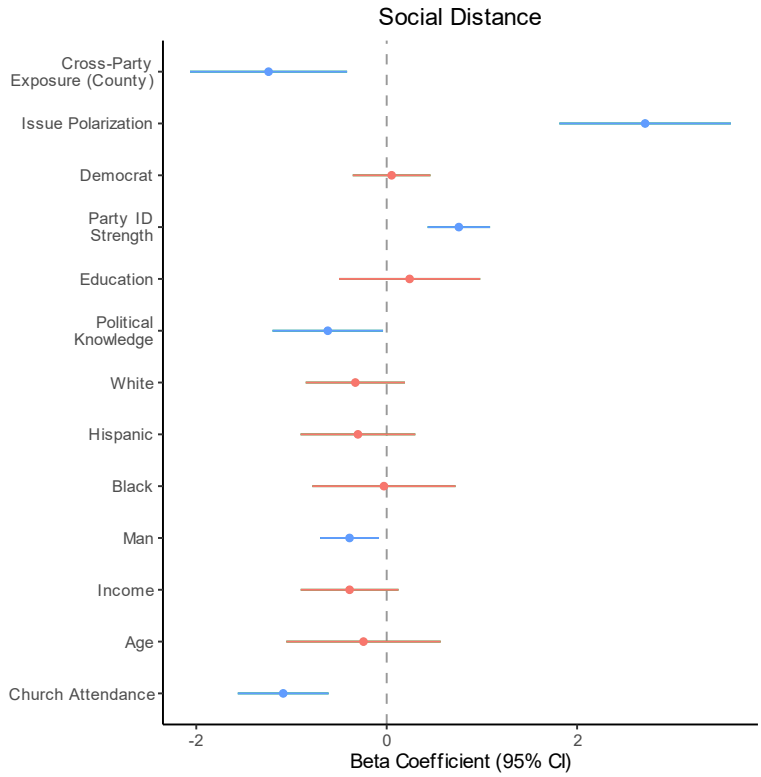
*Predicting affective polarization using precinct-level cross-party exposure in Study 2.*




Next, we examined whether cross-party exposure predicted social distance. In line with our hypothesis, we found a significant negative relationship between cross-party exposure at the *county*-level and social distance, $\beta = -1.24$, $p = .003$, 95% CI = [−2.07, −.42] (see Figure 10). In other words, we would expect that respondents with very little cross-party exposure in their county to score more than a full point lower on the seven-point social distance scale than respondents with very high levels of cross-party exposure, all else being equal. Likewise, we found support for our hypothesis using *precinct*-level cross-party exposure as well. At the precinct level, there was a significant negative relationship between cross-party exposure and social distance, $\beta = -.76$, $p = .04$, 95% CI = [−1.48, −.04] (see Fig. 11).

**Figure 10**

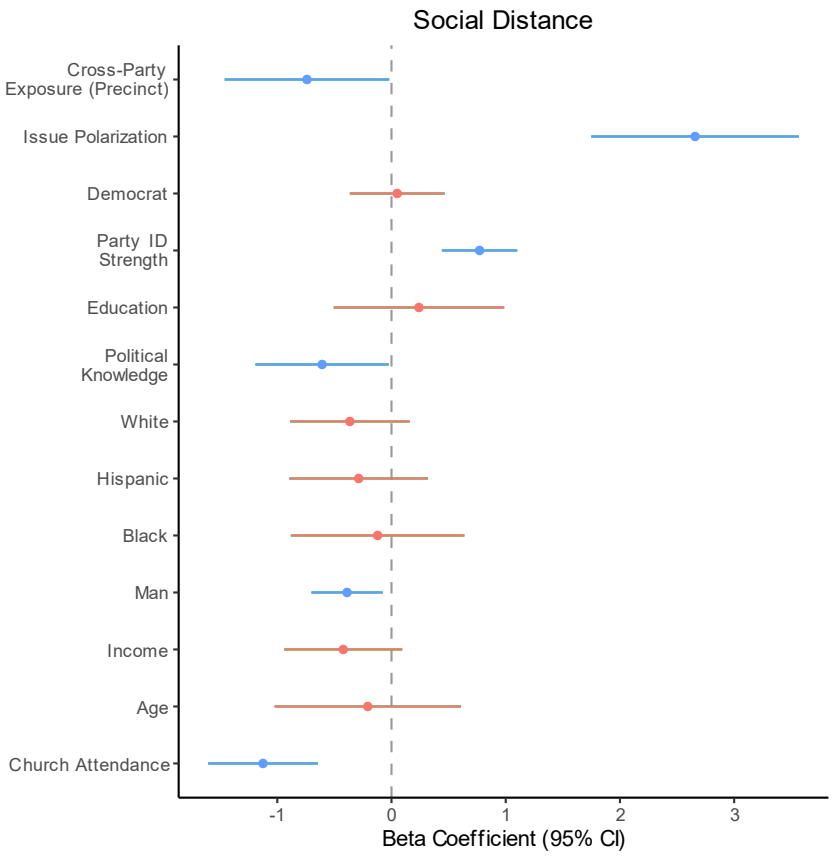
*Predicting social distance using county-level cross-party exposure in Study 2.*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

**Figure 11**

*Predicting social distance using precinct-level cross-party exposure in Study 2.*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

### *Exploratory (Not Pre-Registered) Analyses*

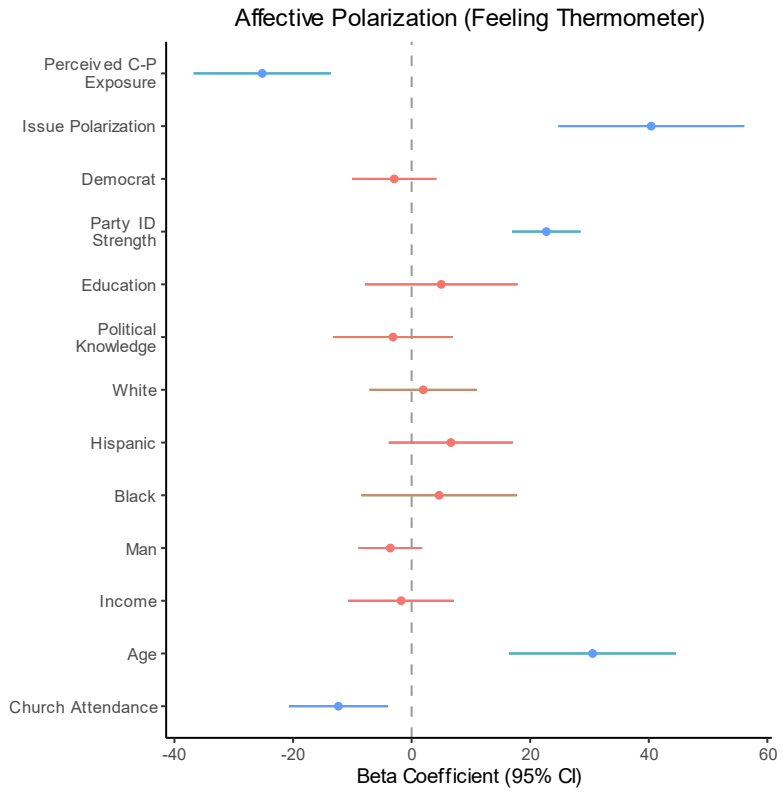
One variable of note that predicted lower levels of affective polarization and social distance was the frequency of church attendance. Church attendance was not a predictor of party-directed affective polarization in Study 1, but it was a meaningful predictor of lower levels of voter-directed affective polarization in Study 2. Because previous research has suggested that overlapping Christian and Republican identities can heighten affective polarization (e.g., Mason, 2018), we ran the same models for just Republican respondents to see if this effect held true for

them as well. Although the effects were smaller than in the sample as a whole (suggesting that church attendance was at least as strong of a predictor for Democrats), we found that more frequent church attendance predicted significantly lower levels of affective polarization (β = −14.09, $p$ = .008, 95% CI = [−24.58, −3.60]) and social distance (β = −.74, $p$ = .01, 95% CI = [−1.31, −.17]) for Republican respondents (see Appendix S2 for full models).

In addition to our more objective measures of cross-party exposure within communities (precinct-level and county-level voting), respondents also indicated the proportion of Democrats or Republicans they thought they interacted with in their communities on a regular basis. We examined how perceived contact predicted voter-directed affective polarization and social distance, using the same covariates as in the previously described pre-registered models. We found that perceived cross-party contact significantly predicted lower levels of voter-directed affective polarization (β = −25.2, $p$ < .001, 95% CI = [−36.76, −13.64], see Fig. 12) and in a separate model it also significantly predicted lower levels of social distance (β = −1.68, $p$ < .001, 95% CI = [−2.33, −1.03], see Appendix S2).

**Figure 12**

*Predicting Voter-Directed Affective Polarization using Perceived Cross-Party Contact in Study 2.*



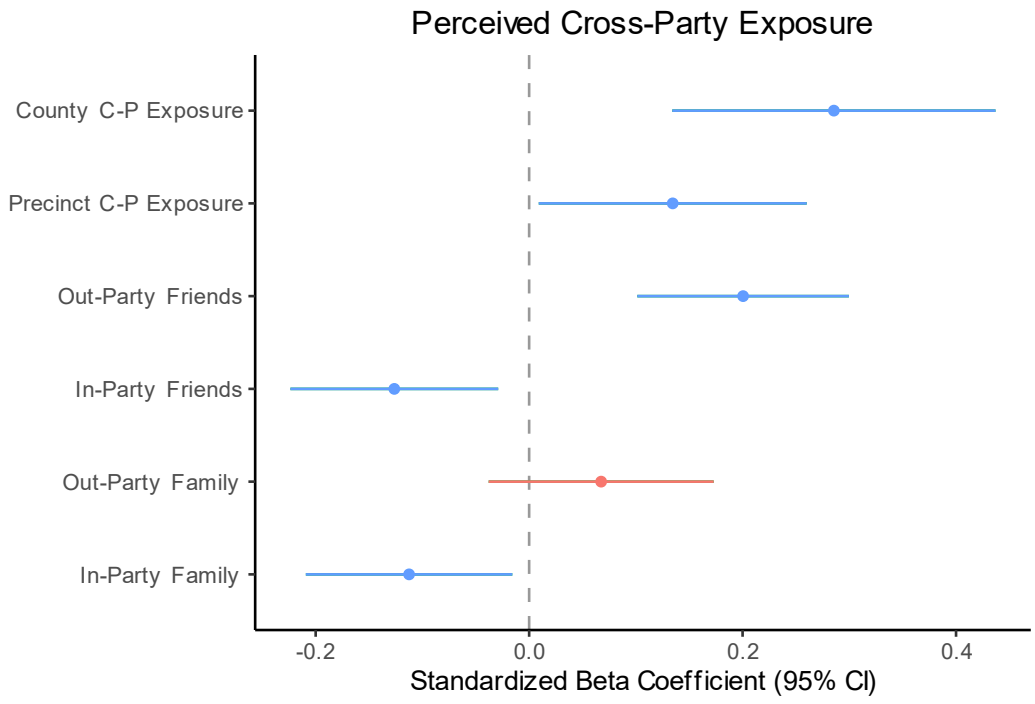

*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant. "C-P" = "cross-party."

Because this measure of perceived cross-party contact is a strong predictor of both voter-directed affective polarization and social distance, and because it is subjective in nature, we examined how it relates to other measures of partisan exposure and contact. We ran a linear regression predicting perceived cross-party contact based on county and precinct level cross-party exposure, in-party and out-party friendships, and in-party and out-party family members (see Fig. 13). We found that perceived cross-party contact was significantly predicted by five of the six variables, suggesting that respondents were likely taking into account many facets of their lives when evaluating the proportion of Democrats and Republicans they interact with in their

everyday lives. More cross-party exposure at the county (β = .29, $p$ < .001, 95% CI = [.13, .44]) and precinct levels (β = .13, $p$ = .04, 95% CI = [.01, .26]) predicted more perceived cross-party exposure, as did having more out-party friends (β = .20, $p$ < .001, 95% CI = [.10, .30]). Having more in-party friends (β = −.13, $p$ = .01, 95% CI = [−.22, −.02]) and family members (β = −.11, $p$ = .02, 95% CI = [−.21, .02]) predicted lower levels of cross-party exposure. Out-party family members was the lone non-significant predictor in this model (β = .07, $p$ = .21, 95% CI = [−.04, .17]).

**Figure 13**

*Predictors of perceived cross-party exposure in Study 2.*

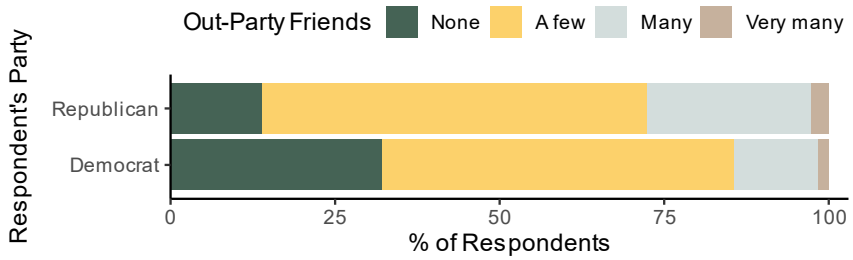


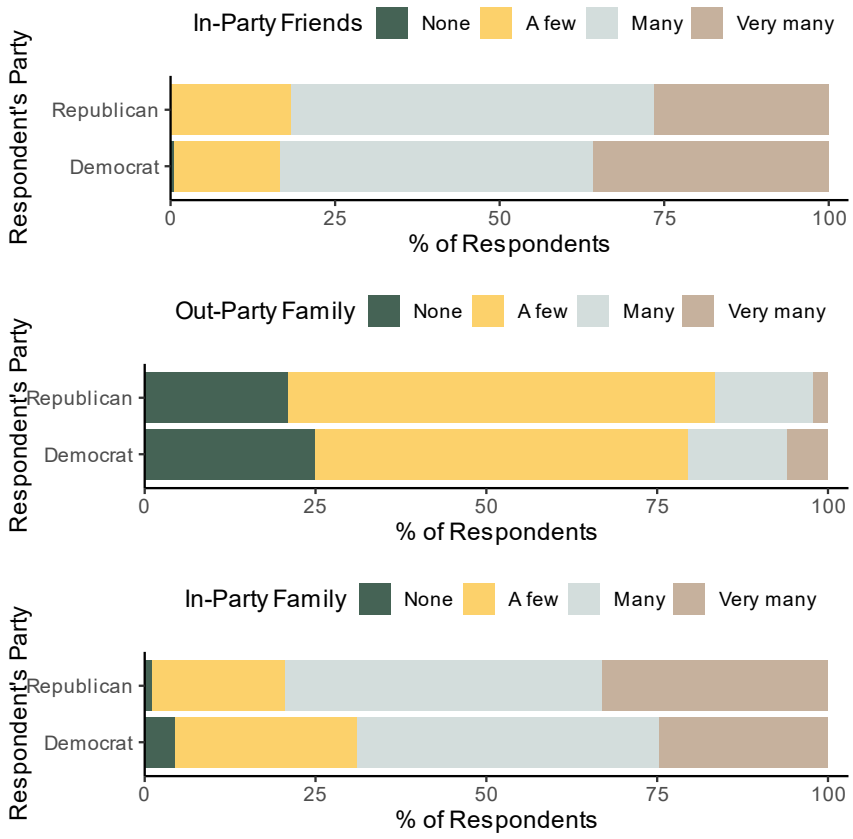
*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant. "C-P" stands for "cross-party."

Participants had many more friends and family members from their in-party than their out-party (see Fig. 14). Respondents frequently reported no friendships with any out-partisans (32.0% of Democrats, 13.8% of Republicans), or were friends with just a few (53.6% of Democrats, 58.6% of Republicans). In contrast, most respondents reported having many friends (47.5% of Democrats, 55.2% of Republicans) or very many friends (35.9% of Democrats, 26.5% of Republicans) from their in-party. Interestingly, this pattern was similar for family members, with many respondents reporting they had no family members from the out-party (24.8% of Democrats, 21.0% of Republicans) and a majority having just a few family members from across the aisle (54.7% of Democrats, 62.4% of Republicans). Most respondents had many (44.2% of Democrats, 46.4% of Republicans) or very many (24.9% of Democrats, 33.1% of Republicans) family members from the in-party.

**Figure 14**

*Out- and In-Party Friendships and Family Members by Party in Study 2*






*Note.* Out-party friendships are shown in the top panel and in-party friendships are shown in the second panel. Out-party family members are shown in the third panel, while in-party family members are shown in the fourth (bottom) panel.

## Discussion

Consistent with our hypotheses, both county and precinct-level measures of cross-party exposure predicted significantly less affective polarization in the form of social distance. These findings suggest that individuals who live in communities with more out-partisans feel more comfortable having interpersonal relationships (coworkers, neighbors, close friends, romantic partners) with members of the opposite party. This pattern emerges even after controlling for the effects of strong predictors of social distance including partisan identity strength and issue

polarization. Furthermore, the predictive power of cross-party exposure at both the county and the precinct levels suggests that the political makeup of communities of various sizes are important for attitudes towards partisans. These results, though correlational, add further support to the theory that cross-party contact may be an important factor in reducing political prejudice. In Study 3, we use experimental methods to further probe the causal relationship between cross-party contact and affective polarization.

Inconsistent with our hypotheses, cross-party exposure did not significantly predict voter-directed affective polarization. This null result held for both county-level and precinct-level measures of cross-party exposure, and for both multiple linear regressions and zero-order correlations. One key difference between Studies 1 and 2 is that in the second study feeling thermometers referred to party voters rather than the parties as a whole. Previous research has shown that affective polarization is generally lower when measuring feelings towards voters rather than parties as a whole (Druckman & Levendusky, 2019). However, it is still surprising the observed relationships between cross-party exposure and affective polarization would be so different between the two studies. This is especially true given that cross-party exposure was a reliable predictor of social distance in Study 2, which is another measure that assesses attitudes related to non-elite partisans.

Unlike county and precinct-level measures of cross-party exposure based on voting patterns, participants' perceived cross-party contact predicted both voter-directed affective polarization and social distance. This measure appears to capture elements of the partisan makeup of counties and precincts in which respondents live, but also their social circles (family and friends). This measure is particularly sensitive to the cross-exposure that people experience, because it takes into account relationships that exist outside of their neighborhood or county of

residence, and also is the result of choices and preferences that individuals make about with whom they want to spend their time. There are likely other elements that affect the degree of contact with out-partisans, such as interactions in the workplace, in schools, in places of worship, and while engaging in leisure activities that may correlate with partisanship. As a result, this measure is likely better at tapping into the level of actual *contact* with out-partisans than our measures of exposure based purely on voting patterns.

An unexpected finding from this study was that more frequent church attendance is associated with *lower* levels of affective polarization, even for Republican participants. This is somewhat surprising in light of research on social sorting (e.g., Mason, 2018) which suggests that the overlap between Republican and Christian identities is related to *higher* levels of affective polarization.[12] However, in Mason's research the outcome variable is thermometer evaluations of *parties* as a whole, not directed at party *voters* as we used for this study. In Study 1, using similar models, church attendance was not a significant predictor of party-directed affective polarization. Might church attendance (and Christian identity more broadly) have different effects on feelings towards individual voters than feelings towards parties as a whole? Future research should be conducted to see if this pattern replicates, but this finding appears consistent with tenants of the Christian faith such as Jesus' call to love one's enemies (Matthew 5:43-48).[13]

Another striking finding from this study is just how few friends and family members have from the out-party. We found that many partisans have very few friends across the aisle, if they

[12] We did not measure participants religious identity, but Republicans are overwhelmingly Christian (Pew Research Center, n.d.), so it can be assumed that "church attendance" represents participation in Christian churches for most Republican respondents. According to this study, conducted in 2014, 82% of Republicans (and leaners) identify as Christians, 3% with a non-Christian faith, and 14% are unaffiliated ("religious nones").

[13] This finding is also consistent with the aphorism "hate the sin, love the sinner", under the (not-so-farfetched) assumption that for some Christian Republicans, aspects of Democratic policies and positions can be viewed as sinful.

have any at all. This is perhaps to be expected, as people have a certain degree of choice when it comes to friendships, and people tend to befriend others with more similar political views (Mosleh et al., 2021; Goldenberg et al., 2020) or end friendships with those whom they disagree politically (Cox, 2021). The same cannot be said to the same extent when it comes to family. Though there are elements of a person's family that can be chosen (e.g., who to marry, whether or not to have kids, the choice to stay in contact with a family member) people generally have less of a choice in who their family members are than who their friends are. We find that most people have many more family members from the in-party than the out-party. An extensive literature suggests that parents transmit political partisanship to their children at a relatively high rate (e.g., Jennings & Niemi 1968; Jennings et al., 2009), but there has been much less focus on extended family networks and political partisanship. Taken together, these findings suggest that there are likely many partisans with very few close connections to people from the out-party. In Study 3, we experimentally manipulate close contact between students from politically homogenous universities to examine the causal relationship between cross-party contact and affective polarization.

## Study 3

Studies 1 and 2 used correlational methods to examine the relationship between cross-party exposure within a person's community (county and/or precinct) and affective polarization and social distance. Although cross-party exposure at the county and precinct levels did not predict affective polarization as measured by feeling thermometers in Study 2, taken together the results of Studies 1 and 2 suggest that being in close proximity to out-partisans is associated with more positive evaluations of the out-party. Despite including many covariates in our analyses, we are unable to make a causal claim about the relationship between contact and affective

polarization. To overcome this limitation of the first two studies, we conducted an experiment to examine whether positive cross-party contact can reduce affective polarization.

Study 2 revealed that many partisans have very few friends or family members from the out-party. In order to use a stringent test of the contact hypothesis, we recruited Democrats from the University of Washington (UW) which is a public university with a large proportion of Democratic students (Niche's #28 most liberal college[14]; 9.5:1 Democrat to Republican ratio) and Republicans from two private Christian universities with a high proportion of Republican students: Liberty University (#5 most conservative college; 6:1 Rep. to Dem. ratio) and Biola University (#9 most conservative college; 7.5:1 Rep. to Dem. Ratio). These students are much less likely to have contact with out-partisans than those at more politically heterogeneous universities. As a result, we were able to extend the research of examine the effect of cross-party contact on individuals with relatively low levels of cross-party exposure.

In Study 3, we formed gender-matched dyads over video chat with one Republican and one Democrat. We assigned half of the dyads to engage in a structured conversation (Fast Friends Protocol; Aron et al., 1997) with one another while the other half of dyads completed a control task (writing short essays). We hypothesized that students who were experimentally assigned to have meaningful contact with an out-partisan would report greater decreases in affective polarization and social distance than those in a control group. Santoro and Broockman (2021) conducted a dyadic study where pairs of Democrats and Republicans discussed either their perfect day or discussed politics, and found that discussing their perfect day could reduce affective polarization but talking about politics did not. Although similar, our study involves participants completing the Fast Friends Protocol (Aron et al., 1997) which consists of talking

[14] All Rankings and Democrat:Republican ratios from Niche:
Most Liberal Colleges in America: https://www.niche.com/colleges/search/most-liberal-colleges/
Most Conservative Colleges in America: https://www.niche.com/colleges/search/most-conservative-colleges/

about topics that are both positive (e.g., "What is your most treasured memory?") and negative (e.g., "When did you last cry in front of another person?") as well as encouraging engagement between participants (e.g., "Tell your partner what you like about them; be very honest this time, saying things that you might not say to someone you've just met."). This method has the advantage of encouraging closeness between respondents, while also covering a wide range of topics. Additionally, by using a control condition where the same topics are considered (in essay form), we are able to pinpoint the precise effect of contact on affective polarization. Our study also has participants from politically homogenous universities, where baseline exposure to the out-party is particularly low. Studies 1 and 2 suggest that those who are isolated from out-partisans in their communities could benefit the most from cross-party contact (in terms of reducing affective polarization), making this aspect of our sample a strength for the ecological validity of our investigation.

Additionally, we extend the work of previous studies which have examined the effects of cross-party contact by also testing whether cross-party contact can increase cooperation between Democrats and Republicans on a political task. This is important, as cross-party cooperation is essential for effective legislation and functioning democracy in the United States. We hypothesized that those in the contact condition would be more cooperative on the political task than participants in the control condition.

## Method

### *Participants*

We recruited 45 Democrat-Republican Dyads (90 participants) between October 2020 and June 2022.[15] All dyads were matched on gender. One dyad was excluded from analysis

[15] Data collection is still ongoing. Recruitment has been slower than anticipated, so we are writing up data that we currently have.

because the audio for one of the participants did not work, bringing the total to 44 dyads. All 44 Democrats were recruited from the University of Washington ($M_{Age}$ = 18.95, SD = 1.37; 77% women, 23% men; 41% Asian/Asian-American, 32% White; 20% multiracial). Republicans ($M_{Age}$ = 23.24, SD = 8.52; 77% women, 23% men; 50% White, 16% Asian/Asian-American, 14% Latino/a, 14% multiracial) were recruited from two different universities: Biola University (20 students) and Liberty University (13 residential students, 11 online students). UW and Biola students were given course credit for participating in the study, while Liberty students could choose to receive course credit or a $20 Amazon gift card for their participation.[16] Exactly half of the dyads were in the contact condition, and half in the control condition (22 dyads/44 participants per condition).

***Procedure and Measures***

Both participants in the dyad were let into a Zoom video chat along with the experimenter. Participants were given instructions to access a Qualtrics survey, where they gave their informed consent and then completed a pre-intervention questionnaire (see Supplemental Materials for a full set of questions). Participants were asked two questions about their feelings regarding cross-party interactions. They indicated whether they agree that friendships between Democrats and Republicans are a positive thing, as well as whether they thought that Democrats and Republicans should show each other respect. Next, they indicated their liberalism/conservativism and their political party identification. They also completed measures of partisan identity, using questions such as "When talking about [political party], how often do you use 'we' instead of 'them'?" Next, they completed measures of affective polarization and social distance. They completed measures of affective polarization using feeling thermometers.

[16] Initially, we recruited both Democrats and Republicans from the University of Washington. However, due to the low number of Republicans at UW, we discontinued this study (see Appendix S3 for UW only results).

They were asked to evaluate how warm (or cold) they feel towards the Democratic and Republican parties, as well as towards *voters* from both parties. Then, they indicated how comfortable they are having an out-partisan as a close friend, romantic partner, and a neighbor.

Next, all participants were told that they were paired with their partner because one of them was a Democrat and the other was a Republican. Dyads then completed one of two (randomly assigned) interventions. Those assigned to the contact condition, completed the Fast Friends Protocol (Aron et al., 1997) with one another. Fast Friends is designed to increase closeness between two people. It consists of two people asking each other a series of questions from three sets of questions, with the level of self-disclosure required increasing from set to set. Questions vary in affect, discussing both positive (e.g., "What is your most treasured memory?") and negative topics (e.g., "Of all the people in your family, whose death would you find most disturbing? Why?". Additionally, there are questions/prompts specifically designed to cause partners to observe one another, such as "Tell your partner something you like about them already." For the sake of time, each set was 10 minutes long (compared to 15 minutes used by Aron and colleagues).

Dyads in the control condition independently wrote essay responses to the questions from the Fast Friends Protocol (Aron et al., 1997), rather than discussing them with their partner. This was done to keep constant the content considered as well as the amount of time spent on the intervention. Questions/prompts that involve observing things about their partner were removed for the control condition. Participants were told that their essay responses would not need to be given to the experimenter (they completed them either on a word processor or hand wrote them), but they were required to leave their cameras on during this task so that the experimenter could make sure they were on task.

After the intervention, participants from both conditions completed a post-intervention questionnaire. This questionnaire consisted of the same feeling thermometers and social distance measures completed prior to the intervention.

Upon completion of the post-intervention questionnaire, participants completed a simulated legislation game (SLG). This task was video recorded, so that in the future we can code these interactions. The SLG was designed to examine cooperation between participants in a political task, while attempting to simulate some of the incentives that legislators face. Participants were told that they would have up to 10 minutes to select four proposed legal changes from a list of options (e.g., “free college tuition for all”, “increase restrictions on abortion”, etc.) that the two of them can agree to. Legal changes were listed in two columns, with one for “Democrat Supported” changes and the other for “Republican Supported.” They were also told that at the end of the academic term, a group of Democrats and a group of Republicans review each set of legal changes agreed to by each dyad. The Democrat in the dyad whose legal changes Democrats deemed as most desirable would win a $20 gift card, and likewise for the Republican in the dyad whose legal changes were deemed to be the best by Republicans. As a result, both participants were incentivized to push for legal changes that would be seen as beneficial for other members of their party. If dyads successfully agreed to four legal changes, the task was ended. If they did not reach an agreement in 10 minutes, the experimenter ended the task then.

After the SLG, participants completed a responded to a series of questions about the SLG, such as how cooperative their partner was and how cooperative they were. They also responded to open-ended questions about what was frustrating or surprising about the task. They then completed a series of demographics questions and were debriefed by the experimenter.

## Results

Pre-intervention feeling thermometers and social distance measures suggest that participants from different parties and universities varied substantially in their levels of initial affective polarization. Democrats in our sample displayed high levels of initial voter-directed affective polarization towards voters, with mean differences between in-party and out-party evaluations over 45 points, while Republicans had mean differences lower than 15 (see Table 1). When evaluating parties as a whole, Republican participants demonstrated higher levels of party-directed affective polarization relative to their levels of voter-directed affective polarization (in line with findings from Druckman & Levendusky, 2019). However, Democrats had similar levels of affective polarization with regards to voters or parties as a whole. Additionally, Democrats had mean levels of social distance near the midpoint, suggesting neither comfort nor discomfort with having interpersonal relationships with Republicans. In contrast, Republican had much lower levels of initial social distance, indicating that most were comfortable being friends, neighbors, and even romantic partners with Democrats.

Democrats and Republican participants were similar to one another in their levels of partisan identity strength. Responses to four questions about party ID were scaled (between 0 and 1) and averaged for each participant. Democrats and Republicans had similar levels of identity strength (see Table 1). Additionally, although there were high levels of affective polarization, on average participants agreed that it is important for Democrats and Republicans to show one another respect (all group means over 6 on a 7-point scale), and generally agreed that cross-party friendships are important (with some variation between groups).

**Table 1**

*Pre-Intervention Means by Party in Study 3*

| Measure | Democrats (UW) | Republicans (Biola/Liberty) |
|---|---|---|
| Voter-Directed Affective Polarization (Feeling Therm.) | 46.63 (24.80) | 13.59 (20.56) |
| Party-Directed Affective Polarization (Feeling Therm.) | 47.50 (24.65) | 30.36 (26.63) |
| Social Distance (1-7) | 4.22 (1.50) | 2.67 (1.24) |
| Partisan ID Index (0-1) | .61 (.19) | .56 (.20) |
| Cross-Party Respect Important (1-7) | 6.20 (.79) | 6.86 (.55) |
| Cross-Party Friendships Positive (1-7) | 5.57 (1.13) | 6.54 (.70) |
| N | 44 | 44 |

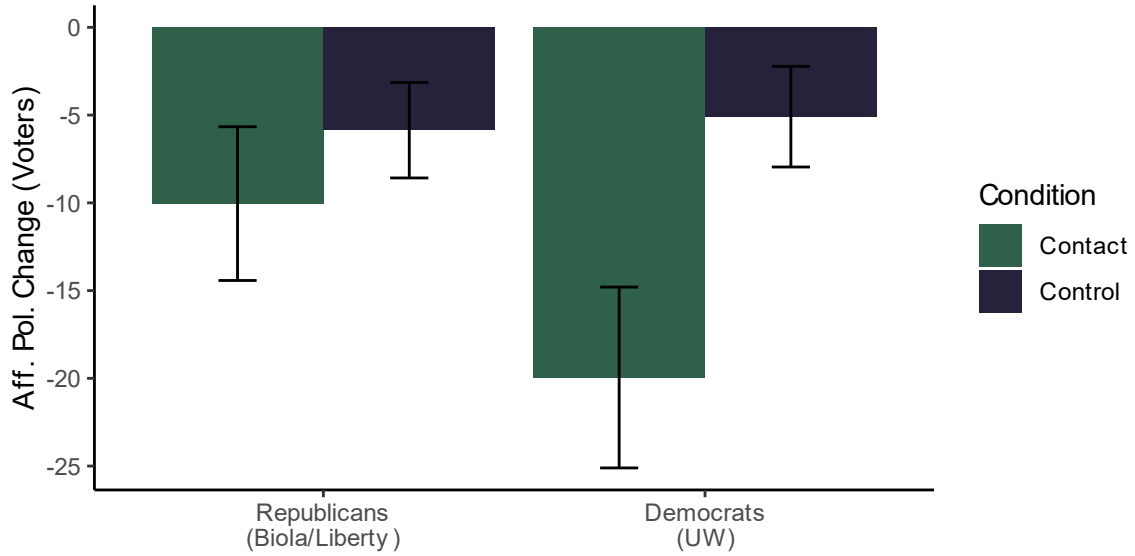
*Note.* SDs in parentheses.

To examine the effect of contact on affective polarization and social distance, we computed change scores for each individual by subtracting post-intervention evaluations from pre-intervention evaluations. We then conducted independent samples t-tests to compare between the contact and control conditions. Voter-directed affective polarization declined significantly more in the in the contact condition ($M = -15.00$, $SD = 22.73$) than in the control condition ($M = -5.48$, $SD = 12.96$), $t(68.3) = 2.41$, $p = .02$, $d = .51$, 95% CI = [.09, .94] (see Fig. 15). This suggests that meaningful positive contact (Fast Friends Protocol) with a cross-party partner reduced partisan bias towards voters more than writing essay responses to the same question independently. The same pattern was observed for a change in party-directed affective polarization, with greater decreases for those in the contact condition ($M = -12.51$, $SD = 19.48$)

than the control condition ($M$ = −3.84, $SD$ = 15.91), $t$(81.0) = 2.27, $p$ = .03, $d$ = .49, 95% CI = [.06, .91] (see Fig. 16). Additionally, social distance decreased more in the contact condition ($M$ = −.30, $SD$ = .70) than for participants in the control condition ($M$ = .09, $SD$ = .96), $t$(78.74) = 2.19, $p$ = .03, $d$ = .47, 95% CI = [.04, .89] (see Fig. 17). In other words, those in the contact condition (vs. control) became more comfortable with having interpersonal relationships with members of the out-party.

**Figure 15**

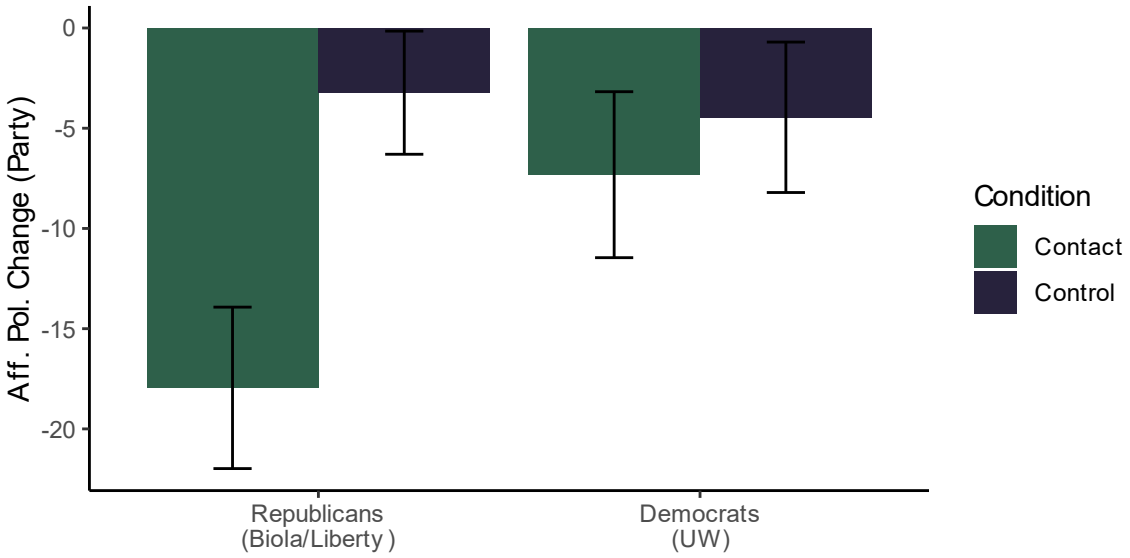
*Voter-Directed Affective Polarization Change by Party and Condition in Study 3*




*Note.* Error bars denote SE.

**Figure 16**

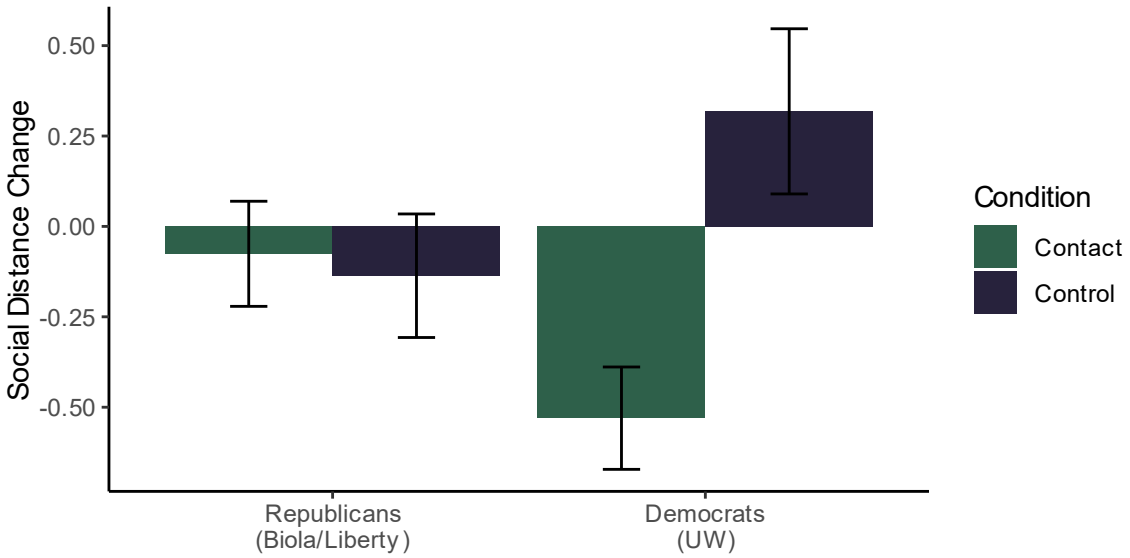
*Party-Directed Affective Polarization Change by Party and Condition in Study 3*




*Note.* Error bars denote SE.

**Figure 17**

*Social Distance Change by Party and Condition in Study 3*




*Note.* Error bars denote SE.

Were there differences between parties in how contact led to changes in affective polarization and social distance? To answer this question, we also fit multiple linear regression models for each dependent variable with dummy coded variables for condition (1 = Contact, 0 = Control) and party (1 = Republican, 0 = Democrat) as well as an interaction term for the two independent variables (Contact X Republican; see Table 2 for full results). Only the linear regression model predicting social distance had a significant interaction term, which suggested that contact led to a greater reduction in social distance for Democrats than Republicans. However, it is important to note that statistical power for these analyses is low.

**Table 2**

*Multiple Regression Models Predicting Affective Polarization and Social Distance in Study 3.*

| Predictor | Voter-Directed Affective Polarization | Party-Directed Affective Polarization | Social Distance |
|---|---|---|---|
| Intercept | −19.96*** [−27.74, −12.17] | −7.32’ [−14.76, .13] | −.53** [−.88, −.18] |
| Contact (Condition) | −14.86** [−25.9, −3.85] | −2.86 [−13.39, 7.67] | −.85*** [−1.34, .36] |
| Republican (Party) | 9.91’ [−1.11, 20.92] | −10.63’ [−21.29, .02] | .45’ [−.04, .95] |
| Contact X Republican | 10.68 [−4.89, 26.26] | −11.86 [−26.84, 3.12] | .91* [.21, 1.61] |
| Multiple $R^2$ | .12 | .10 | .09 |

*Note.* Coefficients are unstandardized. 95% Confidence intervals are displayed in brackets.

***$p$<.001, **$p$<.01, *$p$<.05, ‘$p$<.1.

How cooperative were dyads in the simulated legislation game? Dyads were generally quite successful with regards to agreeing to four legal changes within the ten-minute time frame,

and contrary to our hypothesis there did not appear to be meaningful differences between groups. Eighty-six percent of the (22) dyads in the contact condition were successful and in the control condition 82% of the (22) dyads were successful. Participants evaluated themselves to be very cooperative during the SLG in both the contact ($M$ = 6.43, $SD$ = 1.13) and the control condition ($M$ = 6.34, $SD$ = .89) with no significant difference between conditions, $t(81.47) = .42$, $p = .68$. There were also ceiling effects for how cooperative participants evaluated their partners to be with very high ratings in both the contact ($M$ = 6.61, $SD$ = .81) and control conditions ($M$ = 6.27, $SD$ = 1.11) with no significant differences between groups, $t(78.9) = 1.65$, $p = .10$. There were no significant difference between the number of democratic legal changes agreed upon by those in the contact condition ($M$ = 2.33, $SD$ = .77) compared to the control condition ($M$ = 2.47, $SD$ = .70), $t(34.2) = .58$, $p = .56$.

**Discussion**

Building upon correlational findings in Studies 1 and 2, in Study 3 we find evidence that cross-party contact can *cause* affective polarization to decrease. When students from the Democrat-Republican dyads from politically homogenous universities completed the Fast Friends protocol (Aron et al., 1997), their levels of affective polarization (towards voters and the party) and social distance decreased significantly more than those in the control condition. Importantly, these cross-party interactions consisted of discussing topics ranging from negative life experiences to hopes, dreams, and accomplishments. Nevertheless, the ameliorating influence of intergroup contact appears to be beneficial for reducing intergroup bias in the political domain.

Another finding to note is the difference in the levels of initial (pre-intervention) affective polarization between Democrats and Republicans. This difference was particularly pronounced

when looking at feeling thermometer ratings of party voters, where UW Democrats felt approximately 40 points warmer towards Democratic voters than Republican voters while Biola/Liberty Republicans had in-party preferences of approximately 15 points. Importantly, affective polarization of 40 points as shown by UW Democrats is actually quite similar to national averages for party-directed affective polarization (e.g., Iyengar et al., 2019), and similar to the average levels of voter-directed affective polarization in Study 2. Still, this raises the question: why are the levels of affective polarization so low for Republicans in both versions of the study?

The reason why Biola and Liberty Republicans have relatively low levels of affective polarization is potentially the result of the Christian nature of these universities (Biola's motto is "Above all give glory to God") and the religiosity of their students. In Study 2, more frequent church attendance for Republicans predicted lower levels of voter-directed affective polarization (and also lower social distance). Although we cannot know for certain, it is possible that the effects of religious identities and/or frequent church attendance may be partially responsible for these low levels of voter-directed affective polarization and social distance.

One hopeful finding was that participants across conditions and samples were largely very cooperative in the simulated legislation task. Perhaps everyday Americans are simply more cooperative than those who make it to political office. Still, in future research a more difficult political task needs to be developed in order to better study the factors that influence how effective Democrats and Republicans are in cooperating with one another. This is important to investigate, as bipartisan cooperation is essential for effective democracy in the United States.

**General Discussion**

Across three studies, we examined the role that cross-party exposure and contact play in affective polarization. In Studies 1 and 2, living in areas with more exposure to members of the out-party was associated with lower levels of party-directed affective polarization and social distance. Cross-party exposure at the county-level predicted *party*-directed affective polarization, while greater cross-party exposure measured in three different ways (county-level, precinct-level, and perceived exposure) predicted lower levels of social distance. However, contrary to our hypotheses, cross-party exposure at the precinct and county levels did not significantly predict *voter*-directed affective polarization. Yet, greater perceived cross-party contact was a significant predictor of voter-directed affective polarization, suggesting that interacting with members of the opposite party may reduce this form of affective polarization as well.

We followed these correlational studies with an experimental design to test the causal effect of cross-party contact, with a particular eye to how it affects students at politically homogenous universities. We found that when Democratic and Republican students from universities where they had little exposure to the out-party had meaningful contact with one another, that they decreased in affective polarization and social distance. The sample size for this study was not particularly large,[17] but it does provide further evidence for the ameliorating effect of cross-party contact to a growing literature (e.g., Levendusky & Stecula, 2021; Santoro & Broockman, 2021; Rossiter, 2020). In contrast to previous experiments, we specifically recruited participants who were living and learning within partisan bubbles, demonstrating that cross-party contact can reduce bias for those in politically homogenous communities. Taken together, these three studies suggest that the political composition of the places that they live are related to how

[17] A power analysis conducted using the R package "pwr" suggest that we have a power of .64 to detect a moderate effect (Cohen's $d$ = .5) with a two-sample t-test with our current sample (44 per condition) and alpha = .05.

they think about partisans, and that cross-party contact is likely one of the driving factors of this phenomenon.

Communities within the United States have become increasingly politically homogenous over the last several decades (Bishop & Cushing, 2008; see also Bishop 2020), while affective polarization has also increased during this period (Iyengar et al., 2019). Our research demonstrates that geographic sorting is *directly* linked to affective polarization. By examining how individuals' cross-party exposure is related to their evaluations of the out-party and its voters, we show that where people live matters for how they think about political parties.

Although our research program highlights the role of contact in reducing affective polarization and we control for party identity strength and policy views in our correlational studies, it could be posited that affective polarization causes geographic sorting. However, Mummolo and Nall (2017) cast doubt on the idea that the political composition of communities is an important factor for where people move. They note that while research has shown that Democrats and Republicans often differ in community preferences (Bishop & Cushing, 2008), people do not actually move to places with a greater proportion of in-party members. Rather, those living in politically homogenous communities often move someplace *less* homogenous. They find that even those who would have the most interest in moving to a community with more in-party members (e.g., those with strong ideologies) and those with the most ability to move (e.g., the young, the rich) do not usually end up doing so. Instead, other factors such as cost of living, job opportunities, and safety are much more important considerations which dwarf the importance of a community's partisanship. Based upon this research and experimental findings (including our own) that suggest that more cross-party contact can reduce affective

polarization, we have good reason to believe that exposure to the out-party can lead to less partisan bias.

Our research highlights the importance of the subtle differences in how affective polarization is measured and suggests that perhaps these measures are more different than previously thought. Druckman & Levendusky (2019) found that evaluations of out-parties as a whole (and party elites) are generally more negative than feelings towards out-party voters. We observed a similar pattern for pre-intervention affective polarization in Study 3, where voter-directed affective polarization (but not party-directed affective polarization) was surprisingly low for Republicans students in our sample. We also found that cross-party exposure within one's community predicted significantly less party-directed affective polarization (Study 1) and social distance (Study 2), but not voter-directed affective polarization (Study 2).[18]

Though not the main focus of our investigation, we also observed a curious pattern for the relationship between church attendance and different measures of affective polarization and social distance. Church attendance was not a significant predictor of party-directed affective polarization (Study 1), but more frequent church attendance did predict lower levels of voter-directed affective polarization and social distance (Study 2). Remarkably, this pattern held even when only including Republican participants. This is surprising given work on social sorting that suggests that stronger Christian identities among Republicans is associated with higher levels of affective polarization, rather than less (Mason, 2018). Importantly, Mason's work relies upon measures of party-directed affective polarization, rather than voter-directed affective polarization or social distance. We also saw abnormally low levels of (pre-intervention) voter-directed affective polarization and social distance, for Republican students from Biola University and Liberty University, two schools where Christianity plays a central role. Interestingly, party-

[18] However, perceived cross-party contact did predict voter-directed affective polarization (not pre- registered).

directed affective polarization for these students was much closer to overall averages. Might there be something unique about attending church or being a Christian that affects party vs. voter evaluations in different ways? It is conceivable that many Christians are applying the adage to "hate the sin, love the sinner" when thinking about political parties and their supporters, although future research is certainly needed to investigate this possibility.

**Limitations, Constraints on Generality (COGs), and Future Directions**

As discussed throughout this article, there are several limitations to the present research. The first two studies are correlational in nature, so there is the potential for bidirectionality of the effects, as well as unmeasured confounding variables. There are likely other meaningful factors that covary with partisanship in different localities, such as average income, population density, walkability, racial/ethnic diversity, etc. Future research should seek to examine how these aspects of place relate to affective polarization. Additionally, as we identified, even slight changes in how affective polarization is measured can lead to meaningful differences in the conclusions we draw. Future research is needed to further examine the different ways in which this phenomenon can be measured, and what the differences between them reveal.

As noted previously, our experimental design had a relatively small sample size. Additionally, while one of the strengths of the study was that we recruited participants from three politically homogenous universities, it is still possible that there is something unique about these particular institutions. As a result, care must be taken when generalizing from these findings. Future research should seek to examine how affective polarization manifests in other universities, as well as whether cross-party contact can reduce affective polarization for non-students who live in politically homogenous areas.

**Conclusion**

Our research suggests that where a person lives affects how they think about political parties and those who identify with them. More specifically, we found that when people live amongst larger proportions of out-partisans that they tend to have lower levels of party-directed affective polarization and social distance (but not voter-directed affective polarization), even when controlling for factors such as party identity strength and policy views. We suggest that the ameliorating effects of intergroup contact explains why individuals who have more cross-party exposure in their communities generally have more favorable views of the out-party. By using experimental methods, we demonstrated that cross-party contact reduces affective polarization for individuals attending politically homogenous universities.

# ARTICLE 2

Family Ties: Perceived College Student-Parent Political Differences and Relationship Quality

Adam H. Smiley & Cheryl R. Kaiser

Department of Psychology, University of Washington

Author Note

This project is supported in part by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP).

Correspondence concerning this article should be addressed to Adam H. Smiley, University of Washington, 119A Guthrie Hall Box 351525, Seattle, WA 98195-1525.

Email: ahsmiley@uw.edu

Abstract

Politics cause friction in many interpersonal relationships as political polarization continues to increase in the United States – but does this divisive force harm family relationships as well? In three studies, we examine how perceived political difference relates to three distinct measures of relationship quality for college students and their parents. We find limited evidence that perceived political difference is associated with meaningfully worse relationship quality for college students and their parents when we control for other factors such as non-political differences. The few student perceptions of politics-related relationship impairments we did observe were statistically small, and there was no evidence that parents' perceptions of political-related differences harmed their relationships with their children. This provides initial evidence that the parent-child relationship may be one that is largely immune from the harmful effects of political polarization.

> if you're in michigan and 18+ pls for the love of god do not vote for my dad for state rep. tell everyone
>
> -Stephanie Regan (@streeganz) via Twitter, June 23rd, 2020

Robert Regan, a Republican candidate for the 2020 Michigan House of Representatives, had deep political disagreements with his daughter Stephanie Regan (Folley, 2020). They also held differing views on the extent of white privilege and systemic racism. As a result, Stephanie tweeted that Michiganders should *not* vote for her father. The tweet received nationwide coverage and Robert Regan would go on to lose the Republican primary, finishing third. Though political differences between young adults and their parents rarely go viral, these disagreements are not uncommon (Jennings et al., 2009; Jennings & Niemi, 1968). However, little is known about how political differences affect the parent-child relationship. Are political differences normally harmful to these relationships, or are most political disagreements relatively harmless in the context of a generally stable parent-child relationship? In this article, we investigate the association between perceived political differences and the relationship quality of American college students and their parents.

**Parent-Child Bond**

College students are more independent than children and adolescents, yet most still depend on their parents psychologically and practically. Parental support is associated with higher self-esteem for young adults (Noller & Callan, 1991) and positive relationships with parents into adulthood is associated with lower psychological distress for men (Barnett et al., 1992). College students who received support from parents in striving for shared goals were more motivated and achieved more academically (Kriegbaum et al., 2016). In 2021, 36% of young men (ages 18-29) and 24% of young women indicated that they talk to their parents about

personal problems before they talk to anyone else (Cox, 2021). Young adults often rely on their parents to help in an emergency (Cooney & Uhlenberg, 1992) or to talk through difficult life decisions (Pizzolato & Hicklen, 2011; Cooney & Uhlenberg, 1992). Parent instruction about money can reduce the likelihood that college students will accrue large amounts of credit card debt (Novilitis & MacLean, 2010), and many young adults receive financial assistance from their parents (Martin, 2019; Dickler, 2018; Grant, 2017). Describing young adults' increasing reliance on their parents, Ambert (1997) wrote that "the 'empty nest' syndrome of earlier decades is replaced by a 'cluttered nest' or 'revolving door' family" (p. 71). Since then, young adults have become even more dependent on their parents. In July of 2020, the majority (52%) of young adults aged 18 to 29 lived with their parents (Fry et al., 2020).

This relationship is important for parents as well. A strained relationship with their adult offspring can result in higher levels of psychological distress for parents (Umberson, 1992). Many aging parents require practical assistance, and their adult children often take on some of these caretaking duties, particularly daughters (Schulz et al., 2016; Ambert, 1997). Between 1970 and 2017, the age at first birth for mothers increased from 21.4 years to 26.8 years (Guzzo & Payne, 2018). As a result, parents of college students are now older than they were 50 years ago, meaning that some students may already be caring for their aging parents or will be soon. The mutual reliance between young adults and their parents bolsters their relationship and acts as an incentive to maintain a harmonious relationship.

When the parent-child relationship functions properly, parents offer a "secure base" to their children, giving them a sense of security and facilitating their exploration of their environment (Bowlby, 1988). Even amidst the turbulent adolescent years, the parent-child bond usually remains strong and is the amongst the most stable and permanent relationships (Ambert,

1997; Laursen, 1993). Indeed, most conflicts between adolescents and their parents do not harm the relationship; family members generally tolerate minor quarrels that might be harmful to other forms of relationships (Laursen, 1993).

**Partisan Socialization and Parent-Child Political Disagreement**

Family members play a prominent role in the development of an individuals' views on politics (Hyman, 1959). Parents transmit their partisan affiliations to their children at a high rate (e.g., Jennings & Niemi, 1968; Niemi et al., 1978; Jennings et al., 2009), and transmission has increased meaningfully in recent years (Iyengar et al., 2018). However, parents and their children are less likely to agree on specific policy and social issues (Jennings & Niemi, 1968; Niemi et al., 1978; see also Degner & Dalege, 2013). Attending college itself and interacting with other students can also lead to political attitude change (Newcomb, 1943; Milem, 1998) potentially leading to further political differences between students and their parents. As a result, most college students have at least some political disagreements with their parents. The parent-child relationship is particularly resilient, but is the rising tide of political prejudice in the United States a threat to this important bond?

**Affective Polarization**

Affective polarization, defined as *hostility* and *social distance* between people of different political parties, has been on the rise in the United States for decades (Mason, 2018; Iyengar et al., 2019). Hostility is often measured as an individual's evaluation of their own political party (on a feeling thermometer from 0-100 degrees) minus their evaluation of an opposing political party.[19] According to American National Election Study (ANES), the difference between in-party and out-party evaluations has increased dramatically in the past forty years, rising from 23 degrees in 1978 to 41 degrees in 2016 (Iyengar et al., 2019). This rise in the

[19] In the United States, Democrats and Republicans are used almost exclusively.

United States has outpaced that in other developed countries over the last four decades (Boxell et al., 2020). Among Americans, implicit bias against out-party members is now even stronger than implicit bias towards members of racial out-groups (Iyengar & Westwood, 2015).

With political animosity at a fever-pitch, it is unsurprising that social distance between partisans is also substantial. The 2021 American Perspectives Survey found that 65% of Democrats and 44% of Republicans have just a few friends or no friends at all from the opposing political party (Cox, 2021; see also Smiley & Kaiser, in prep.). Political views can cause existing friendships to end and prevent new ones from forming. Twenty-eight percent of Democrats and 10% of Republicans reported ending friendships due to political differences, with 22% of those friendships ending due to disagreements about President Trump (Cox, 2021). Mosleh and colleagues (2021) conducted a field experiment on Twitter where they followed Twitter users on the platform using fake accounts that projected various political orientations. They found that users were three times as likely to follow-back accounts with a shared political lean.

One potential explanation for why politics can be particularly harmful to relationships is that they can be framed in moral terms. People have stronger emotional responses when they hold moral convictions about an issue (Skitka et al., 2005), and intolerance towards those who hold different attitudes is higher when a topic is viewed as moral (Skitka, 2010). Additionally, adolescents and their parents can use small disputes to quarrel over much deeper issues (Noller & Callan, 1991), and it is possible that politics could be used to argue about underlying principles of right and wrong. In our investigation, we examine how attributing political disagreements to differences in morality is related to relationship quality.

How do political disagreements affect family relationships? A large majority (85%) of 2020 ANES respondents indicated that political differences had either not harmed or only

harmed their family relationships a little bit in the previous four years (Abrams, 2021). However, politics can be a force of division within families. In 2016, Thanksgiving dinners were 30-50 minutes shorter when guests were visiting from an area with a different political lean (vs. visiting from an area with the same political lean), with the effect being particularly strong in areas where political advertising was substantial (Chen & Rohla, 2018). Political views affect who individuals want in their family. In 2010, 49% of Republicans and 33% of Democrats indicated they would be upset if their child married someone from the opposite political party (Iyengar et al., 2012). Likewise, people prefer to date those who share their political views (Huber & Malhotra, 2017). This pattern extends to marriage, with 81.5% of American spousal pairs identifying with the same political party (Iyengar et al., 2018). While these studies suggest that political differences can drive a wedge between family members, our research is the first to examine how political difference is related to relationship quality of parents and their children. As the parent-child bond is uniquely strong, this investigation presents a stringent test case for the effects of political bias.

In three studies, we investigate how perceived political difference relates to college student-parent relationship quality. In Study 1, we survey college students, and in Studies 2 and 3 we survey college students and their parents.

## Study 1

In Study 1, we examined whether a student's perceived political difference with their parents predicts lower relationship quality. We hypothesized that greater perceived political difference would predict worse relationship quality on three distinct measures: closeness, social support, and negative interactions. Furthermore, we hypothesized that this effect would be

significantly greater than a standardized beta (β) coefficient[20] of .14, using minimum-effect testing (Smiley et al., under review). This value is equivalent to estimates of the "ambient noise level" (Lykken, 1968; also referred to as the "crud estimate," Orben & Lakens, 2019), which is based upon the average correlations found between theoretically unrelated variables (Standing, 1991; Webster & Starbuck, 1988). Additionally, we predicted that attributing political dissimilarity to differences in morality will moderate the effect of political difference on relationship quality. That is, we expected that when political differences are attributed to morality, the relationship between political dissimilarity and relationship quality will be even stronger. The above hypotheses were pre-registered.

**Method**

**Participants.** Two hundred eleven undergraduates (61% women, 38% men, 0.5% non-binary; $M_{AGE}$ = 19.6, *SD* = 2.28; 54% Asian, 26% White, 8% multiracial, 5% Black, 5% Latino/a, 2% Middle Eastern, 0.5% Native Hawaiian/Pacific Islander, 0.5% Native American) from the University of Washington participated in this study for credit in their psychology courses. Students could provide information about their relationship with up to two parents/guardians, and in total students completed surveys about relationships with 405 parents/guardians. Because the study involved asking participants about their relationships with parents, only students who indicated they had contact with at least one parent/guardian were eligible to participate. Data was collected January-March of 2021, shortly after the January 6th Capitol riots and the 2020 US Presidential Election.

**Procedure.** To begin, participants specified the nature of their relationship (e.g., "mother", "step-mother", "uncle", etc.) with up to two parents/guardians with whom they had contact. Next, participants responded to a series of questions with regards to each of the

[20] β is used to denote standardized beta coefficients, throughout.

parents/guardians they specified previously, as well as questions about their relationship with that person. Participants also completed a basic demographics questionnaire as well as several questions for exploratory purposes (see Supplemental Materials for a complete list of materials).

**Measures.** Our pre-registered models included the following variables. Covariates were selected to control for factors unrelated to politics that may affect relationship quality.

***Outcome Variables.***

*Closeness (Inclusion of Other in Self)*. The single-question Inclusion of Other in the Self scale (IOS; Aron et al., 1992) consists of selecting an image of overlapping circles that best represents one's current relationship with their parent/guardian, where one circle was labeled as "self" and the other labeled as "other" (1-7 scale). Higher scores represent more overlap and more interpersonal closeness.

*Support.* The 21-item support subscale from the Social Provisions Version scale (NRI-SPV; Furman & Buhrmester, 1985) measured how much social and instrumental support students and their parents provide to one another. The support subscale asks questions such as "How much does this person teach you how to do things you don't know?" and "How much do you talk to this person about things that you don't want others to know?" (1-5 scale; "little or none" to "the most"). The value for social support is the mean of all responses, higher scores denote more social support.

*Negative Interactions.* The six-item negative interactions subscale from the Social Provisions Version scale (NRI-SPV; Furman & Buhrmester, 1985) measured the degree to which students and their parents argue with or annoy one another. The negative interactions subscale includes questions such as "How often do you and this person disagree and quarrel with each other?" and "How much do you and this person get annoyed with each other's behavior?" (1-5

scale; “little or none” to “the most”). The value for negative interactions is the mean of all responses, higher scores denote more negative interactions.

***Predictor Variables.***

*Political Difference.* Political difference was measured using a single item, which asked how different/similar students’ political views were to their parent/guardian (1-7 scale; “very different” to “very similar”). Values were recoded so that higher scores represented more political difference.

*Moral Attribution.* Participants indicated the extent to which they believed their political differences with their parent/guardian were a result of moral differences (1-7 scale; “not at all” to “completely”).

*Non-Political Difference.* Non-political difference was measured using a single item, which asked how similar/different students they are to their parents in matters not including politics (1-7 scale; “very different” to “very similar”). Values were recoded so that higher scores represented more non-political difference.

*Number of Parents.* Indicates the number of parents the child provided responses for (1 or 2).

*Number of Siblings.* Indicates the number of siblings the participant has.

*Mother.* Dummy code for if the parent is the student’s mother.

*Father*. Dummy code for if the parent is the student’s father.

*Cisgender Man.* Dummy code for if the student identifies as a cisgender man.

*Cisgender Woman.* Dummy code for if the student identifies as a cisgender woman.

*Live With.* Dummy code for if the student is living with their parent during the academic term when the survey was conducted.

*White.* Dummy code if participant is monoracially White.

*Income.* Family's household income within one of five possible ranges.

**Results and Discussion**

On the 7-point Inclusion of Other in the Self scale (Aron et al., 1992), average responses were close to the midpoint with a mean of 4.11 (*SD* = 1.49). Similarly, they were just above the midpoint on the 5-point support scale (Furman & Buhrmester, 1985), with a mean of 3.30 (*SD* = .76). In contrast, average responses on the 5-point negative interactions scale (Furman & Buhrmester, 1985) were quite low, with a mean of just 2.02 (*SD* = .79), suggesting participants had relatively few quarrels with their parents. See Table 1 for mean scores across studies. Zero-order correlations were consistent with hypotheses that students who perceived more political difference from their parents had lower levels of closeness ($r = -.27$, $p < .001$) and support ($r = -.36$, $p < .001$), and higher levels of negative interactions ($r = .19$, $p < .001$) with their parents.

**Table 1**

*Means (and SDs) for three measures of relationship quality in Studies 1-3 by respondent group*

| Variable | Study 1 | Study 2 | | Study 3 | |
|---|---|---|---|---|---|
| | Students | Students | Parents | Students | Parents |
| Inclusion of other in self (1-7 scale) | 4.11 | 4.41 | 4.75 | 4.38 | 4.81 |
| | (1.49) | (1.33) | (1.40) | (1.47) | (1.47) |
| NRI-SPV Support (1-5 scale) | 3.30 | 3.52 | 3.51 | 3.50 | 3.50 |
| | (.76) | (.74) | (.63) | (.74) | (.62) |
| NRI-SPV Negative Interactions (1-5 scale) | 2.02 | 1.94 | 1.77 | 1.93 | 1.68 |
| | (.79) | (.74) | (.73) | (.83) | (.73) |
| N | 211* | 275 | 275 | 250 | 250 |

*Note*. Means (*SD*) for each scale on each study by respondent. *Study 1 had 211 participants who provided data for 405 parent/guardian relationships.

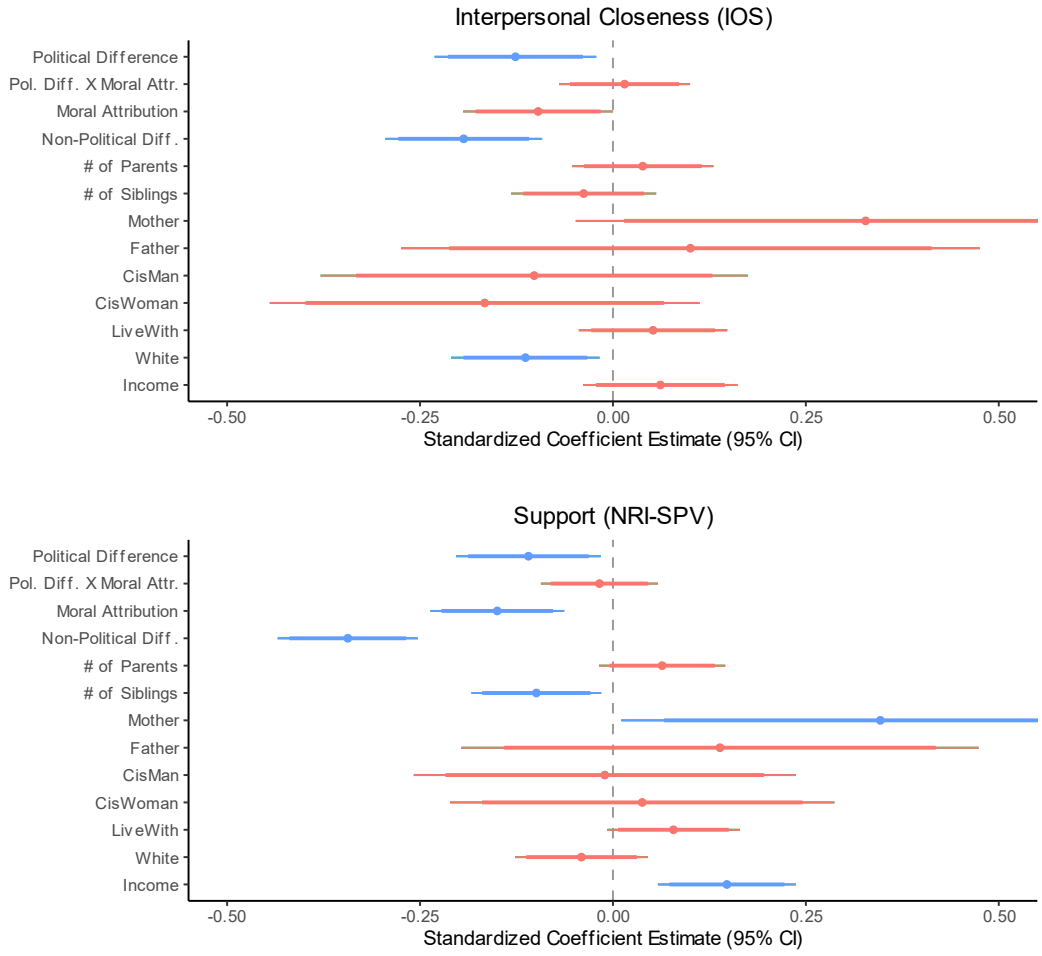
As pre-registered, we conducted multiple linear regressions to test the effect of political similarity on relationship quality while holding other factors constant (see Fig. 1 for a complete list of covariates). Using the principals of minimum effect testing (MET; Smiley et al., under review), we determined that to be considered meaningful, the effect of political difference on relationship quality would need to be significantly more negative than a standardized β of −.14 (or significantly greater than .14 for the negative interactions scale). We found that the relationship between political difference and interpersonal closeness (IOS scale) was significantly more negative than zero, β = −.13, $p$ = .02, 95% CI = [−.23, −.02], showing a conventionally statistically significant effect. But contrary to our more conservative hypothesis, the effect was not significantly more negative than −.14 under MET, 90% CI[21] = [−.21, −.04] (see Fig. 1). Likewise, political difference significantly predicted interpersonal support (NRI-SPV) based on traditional null hypothesis significance testing (NHST), β = −.11, $p$ = .02, 95% CI = [−.20, −.02], but the effect was not significantly stronger than −.14 under MET, 90% CI = [−.19, −.03]. Additionally, the relationship between political difference and negative interpersonal interactions (NRI-SPV) was significantly greater than zero, β = .14, $p$ = .01, 95% CI = [.03, .24], but contrary to our hypotheses it was not significantly greater than .14 using MET, 90% CI = [.05, .22] (see Fig. 1). These results all suggest that political difference is related

[21] MEST uses 90% CIs for an alpha level of .05 (see Smiley et al., under review).

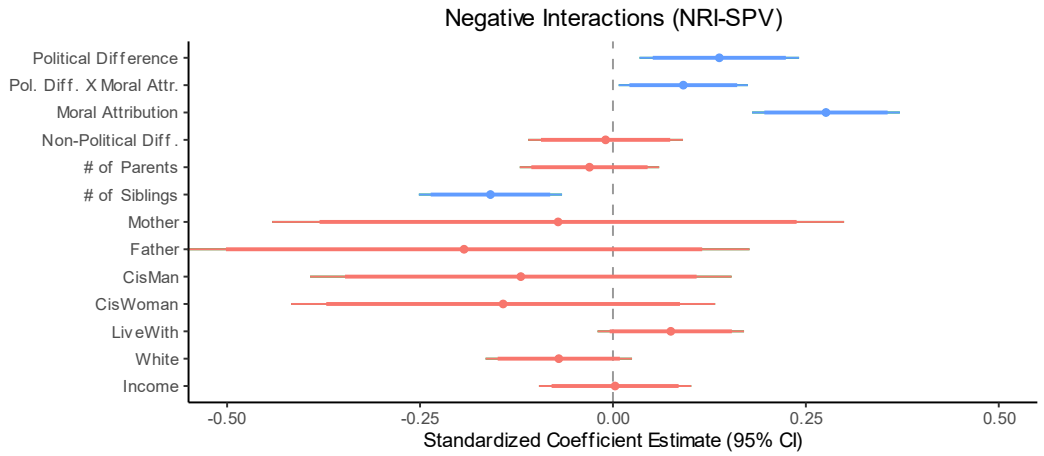
to more negative parent-child relationships, however, the effect was not as strong as hypothesized.[22]

**Figure 1**

*Predictors of Student-Parent Relationship Quality in Study 1*




[22] Equivalence tests for the relationship between perceived political difference and the three relationship quality variables using a null region of −.14 to .14 were all non-significant, indicating that we cannot conclude that the effects were too small to matter.



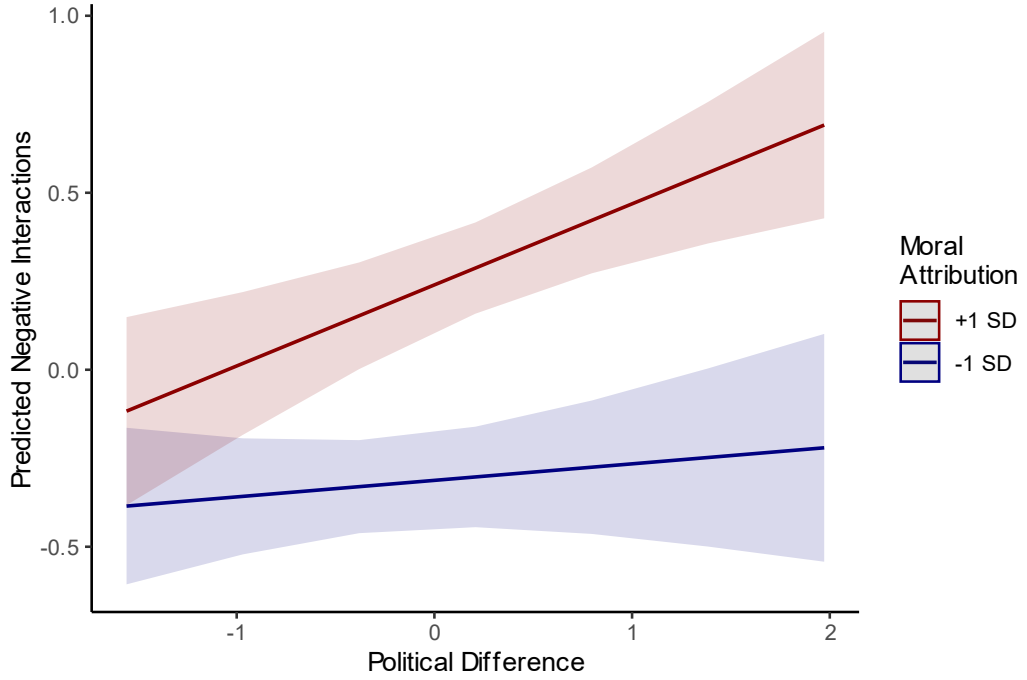

*Note*. Standardized β coefficients from linear regression models predicting three measures of relationship quality. Responses are from students only. Points = estimates. Thick whiskers = 90% confidence intervals. Thin whiskers = 95% confidence intervals. Estimates significantly different than 0 with α = .05 are shown in blue, while non-significant effects are shown in orange.

We also hypothesized that the effect between political difference and the three measures of relationship quality would be moderated by moral attributions to those differences. Contrary to our hypothesis, there was not a significant interaction between political difference and moral attributions when predicting interpersonal closeness (β = .02, $p$ = .72, 95% CI = [−.07, .10]) and support (β = −.02, $p$ = .65, 95% CI = [−.09, .06]). However, there was a significant interaction between political difference and moral attributions when predicting negative interactions between students and their parents, β = .09, $p$ = .03, 95% CI = [.01, .17]. This suggests that political difference was particularly tied to negative interactions when students attributed those differences to moral reasons (see Fig. 2). Simple slopes tests revealed that when respondents were high (+1 *SD*) in moral attributions more political difference was associated with more

negative interactions (β = .23, $p$ = .03, 95% CI = [.10, .36]), while there was no significant relationship between political difference and negative interactions for participants low (−1 *SD*) in moral attributions (β = .05, $p$ = .50, 95% CI = [−.09, .18]).

**Figure 2**

*Moral Attribution as a Moderator of Political Difference in Study 1*




*Note.* Predicted values of negative interactions by political difference and moral attribution (+/− 1 *SD*). Shaded area represents 95% CI. All variables are standardized.

Although not pre-registered, moral attributions for political differences predicted significantly less support between students and parents (β = −.15, $p$ < .001, 95% CI = [−.24, −.06]). Additionally, moral attribution predicted significantly more negative interactions between students and their parents (β = .27, $p$ < .001, 95% CI = [.18, .37]). In fact, moral attributions for

political differences was the strongest predictor of negative interactions between college students and their parents. This highlights the importance of moral values, and how they relate to political views in the context of student-parent relationships. Moral attribution was not a statistically significant predictor of interpersonal closeness, though it was approaching significance ($\beta = -.09$, $p = .05$, 95% CI = [−.19, .00]).

Study 1 found that even after controlling for covariates, perceived political difference predicted worse relationship quality on all three measures (closeness, support, negative interactions). However, the size of the effects was smaller than hypothesized. Study 1 also found mixed results for the role of moral attributions for political difference. Consistent with our hypothesis, moral attributions significantly moderated the relationship between perceived political difference and negative interactions such that political differences were most damaging to interactions when politics were viewed through the lens of morality. However, contrary to our hypotheses, moral attributions did not moderate the relationship between perceived political difference and the positive measures of relationship quality (closeness and support). Significant main effects for moral attribution predicting support and negative interactions further highlights morality as an important aspect for the student-parent relationship. In Study 1 we only collected responses from students, so our perspective was limited to only one half of the relationship. In Study 2, we used similar methods but with a dyadic approach to examine how perceived political difference predicts relationship quality according to both students and parents.

## Study 2

In Study 2, we surveyed college students and one of their parents/guardians about their relationship. Additionally, we included open-ended responses to gain deeper insight into how politics affect the student-parent relationship. We pre-registered our hypothesis that for both

college students and their parents, greater political difference would predict significantly worse relationship quality. Furthermore, we pre-registered that we would consider the effect to be of a meaningful magnitude if it was significantly stronger than a standardized beta of +/− .14 using minimum-effect testing.

## Method

**Participants.** Three hundred six undergraduates (65% women, 33% men, 2% nonbinary; $M_{Age}$ = 18.85, *SD* = 1.92; 64% Asian, 21% White, 8% multiracial, 3% Latino/a, 3% Middle Eastern, 2% Black) from the University of Washington participated in this study for credit in their psychology courses. They each were recruited one of their parents and guardians to participate, and 284 parents/guardians (70% women, 30% men, 0.4% nonbinary; $M_{AGE}$ = 50.02, *SD* = 6.19; 64% Asian, 22% White, 2% Latino/a, 2% Middle Eastern, 2% Black, 2% multiracial[23]) participated in the study. Because the study involved asking participants about their relationships with parents/guardians, only students who indicated they had contact with at least one parent/guardian were eligible to participate. After removing respondents whose parents/children did not complete the survey, we had a total of 275 student-parent/guardian dyads (550 total respondents). Data was collected October-December of 2021.

**Procedure.** To begin, student participants specified the nature of their relationship (e.g., "parent", "step-parent", "guardian", etc.) and the gender identity of one parent/guardian that they have contact with who would also be willing to respond to a survey. Participants then completed the same questions posed to participants in Study 1, with a few additional questions for exploratory analyses.

Student participants were also asked the extent to which they believed their parent/guardian was proud or ashamed of them (both 1-7 scales; "strongly disagree" to "strongly

[23] One parent's self-identified race was "human."

agree"). They were also asked a series of open-ended questions. To examine the extent to which students intentionally avoid political conflict, they were asked "Do you avoid discussing certain political topics with your [parent/guardian]? Why or why not?" To determine whether students practice conflict aversion, they responded to the question "Do you ever end political conversations with your [parent/guardian] if the conversation gets too emotional? Why or why not?" Participants were also asked "Does your [parent/guardian] have any political views that makes you think less of them? If so, why does it make you think less of them?" And finally, they were asked "Does your [parent/guardian] have any political views that you feel are in opposition to one or more of your social identities (i.e., race, gender, sexual orientation, religion, disability, etc.)? If so, how does this make you feel?" Basic demographic information was also collected.

Upon completion of the survey, student participants were provided with a unique survey link to give to their parent/guardian. This link contained a randomly assigned respondent ID number that could be used to link student and parent/guardian responses.

Parents/guardians completed the same survey as the student participants in Study 2, with their responses referencing their relationship with their child. The one exception was that instead of being asked how proud/ashamed they believed their child to be of them, they were asked if they were proud/ashamed of their child.

**Measures.** Measures used in the pre-registered models were the same as those used in Study 1 with a few exceptions. First, rather than having multiple variables measuring parental role (mother and father in Study 1) we used a dummy code for male parents or guardians (1) or not (0). We did the same for student gender, with a dummy code for male students (1) or not (0). Additionally, we did not include a variable for number of parents, as all students completed surveys for only one parent.

Though not specifically included in the pre-registration (due to an oversight), we also added a dummy coded variable for family role of the respondent (Student or Parent) and a political difference X family role interaction term to examine whether the relationship between political difference and relationship quality was different for students than parents. We also did not pre-register that we would include an interaction for perceived political difference X moral attributions, but we included this in our models to be consistent with Study 1.[24]

**Results and Discussion**

Student responses on the closeness and support scales in Study 2 were slightly higher than in Study 1, while the negative interactions ratings were slightly lower than in Study 1 (see Table 1). This is likely a reflection of the fact that in Study 2, students were evaluating a relationship with a parent/guardian who at a minimum would have to be willing to take a survey to help their child earn course credit. In comparison to student evaluations, parents/guardians reported slightly more interpersonal closeness with their child and scored slightly lower on the negative interactions scale than their children did. However, parents' mean support scale scores were quite similar to those from the students. Zero-order correlations were consistent with the hypothesized direction in that those who perceived more political difference from their family member reported lower levels of closeness ($r = -.17$, $p < .001$) and support ($r = -.26$, $p < .001$), and higher levels of negative interactions ($r = .21$, $p < .001$).

As pre-registered, we used multilevel models with random intercepts for each parent-student pair to test the hypothesis that perceived political difference was associated with worse relationship quality on three measures while holding other factors constant. We tested three models (see Fig. 3), one for each measure of relationship quality, which followed our pre-registered analysis plan with one exception (described previously). All variables were

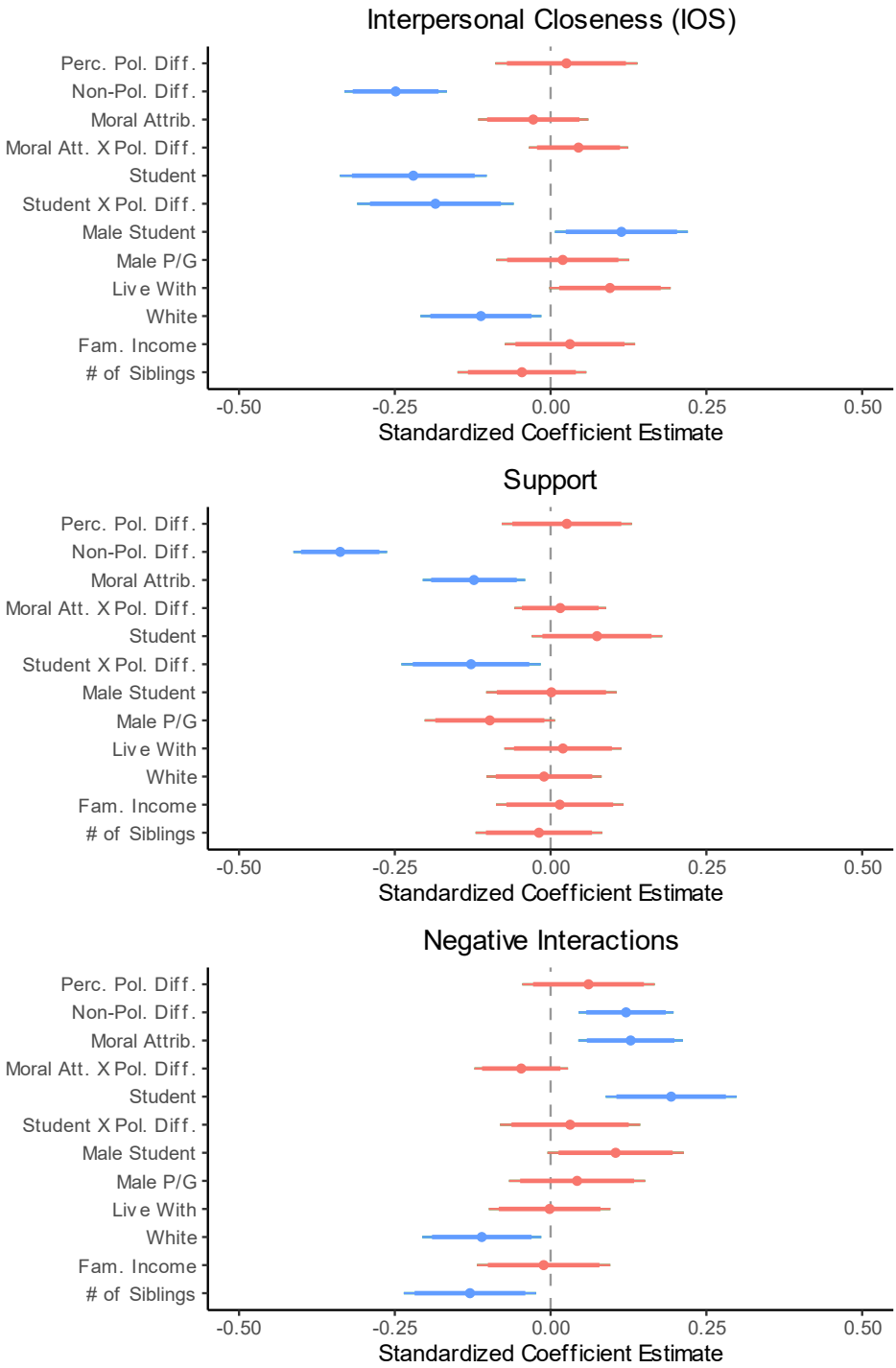
[24] See Appendix S2 for pre-registered models without deviations from the analysis plan.

standardized (mean centered and scaled by standard deviation). Based on a null hypothesis of no effect whatsoever, we determined that perceived political difference did not significantly predict closeness (β = .03, 95% CI = [−.09, .14], $p$ = .66), support (β = .03, 95% CI = [–.08, .13], $p$ = .63), or negative interactions (β = .06, 95% CI = [–.04, .17], $p$ = .26).[25]

**Figure 3**

*Fixed-Effect Predictors of Student-Parent Relationship Quality in Study 2*

[25] As none of these effects were significant with a null hypothesis of no effect whatsoever, all findings were necessarily non-significant by minimum effect testing with H'0 = [–.14, .14]. Additionally, an equivalence test (H'0 = [–.14, .14]) examining the relationship between perceived political difference and interpersonal closeness suggested that the effect was too small to be practically meaningful (β = .03, 90% CI = [–.07, .12]), as did an equivalence test for political difference and support (β = .03, 90% CI = [–.06, .11]). While these findings suggest that main effect of perceived political difference is too small to matter, interactions and simple slopes analyses suggest that this effect is different for students and parents. An equivalence test for the relationship between perceived political difference and negative interactions was not significant (β = .06, 90% CI = [−.03, .15]), suggesting that there was insufficient data to conclude whether the effect is of a meaningful magnitude in the population.


Interpersonal Closeness (IOS)
Perc. Pol. Diff.
Non-Pol. Diff.
Moral Attrib.
Moral Att. X Pol. Diff.
Student
Student X Pol. Diff.
Male Student
Male P/G
Live With
White
Fam. Income
# of Siblings
-0.50
-0.25
0.00
0.25
0.50
Standardized Coefficient Estimate
Support
Perc. Pol. Diff.
Non-Pol. Diff.
Moral Attrib.
Moral Att. X Pol. Diff.
Student
Student X Pol. Diff.
Male Student
Male P/G
Live With
White
Fam. Income
# of Siblings
-0.50
-0.25
0.00
0.25
0.50
Standardized Coefficient Estimate
Negative Interactions
Perc. Pol. Diff.
Non-Pol. Diff.
Moral Attrib.
Moral Att. X Pol. Diff.
Student
Student X Pol. Diff.
Male Student
Male P/G
Live With
White
Fam. Income
# of Siblings
-0.50
-0.25
0.00
0.25
0.50
Standardized Coefficient Estimate

*Note*. Standardized β coefficients of fixed effects for three levels of relationship quality. Models had random intercepts for each student-parent dyad. Responses are from both parents and students. Points = estimates. Thick whiskers = 90% confidence intervals. Thin whiskers = 95% confidence intervals. Estimates significantly different than 0 with α = .05 are shown in blue, while non-significant effects are shown in orange. "P/G" refers to parent or guardian.

To determine if there was a different relationship between political difference and relationship quality for students than there is for parents, we examined the interaction between perceived political difference and family role interaction. In the model predicting interpersonal closeness there was a significant perceived political difference X family role interaction (β = −.18, $p$ = .004, 95% CI = [−.31, −.06]) suggesting that the relationship between perceived political difference and interpersonal closeness is more negative for students than parents. Indeed, a simple slopes analysis revealed that there was a significant negative relationship between perceived political difference and closeness for student respondents (β = −.16, $p$ = .005, 95% CI = [−.27, −.05]) but not for parents (β = .03, $p$ = .66, 95% CI = [−.09, .14]). In the model predicting support, there was a significant perceived political difference X family role interaction (β = −.13, $p$ = .03, 95% CI = [−.24, −.02]) which suggests that the relationship between perceived political difference and support is more negative for college students than their parents. A simple slopes analysis suggested found a marginally significant negative association between perceived political difference and support for student respondents (β = −.10, $p$ = .05, 95% CI = [−.20, .00]) while the relationship was not significant for parents (β = .03, $p$ = .62, 95% CI = [−.08, .13]). And finally, the perceived political difference X family role interaction

was not a significant predictor of negative interactions between students and their parents (β = .03, $p$ = .58, 95% CI = [−.08, .14]). These findings suggest that for positive aspects of relationship quality (closeness and support), there may be an association between more political difference and worse relationship quality for students but not parents.

Though we did not have a pre-registered hypothesis about the relationship between moral attributions to political differences and relationship quality in Study 2, the same pattern emerged as in Study 1. In Study 2, there was again a main effect for moral attributions predicting significantly less support between students and parents (β = −.12, $p$ = .004, 95% CI = [−.20, −.04]), and significantly more negative interactions between students and their parents (β = .13, $p$ = .004, 95% CI = [.05, .21]). Moral attribution was not a statistically significant predictor of interpersonal closeness (β = −.03, $p$ = .54, 95% CI = [−.12, .06]). Additionally, there was not a significant perceived political difference X moral attribution interaction for any measure of relationship quality (see Fig. 3).

See Appendix S2 for full results of shame and pride questions as well as a discussion of patterns in open-ended responses.

In Study 1, we observed that perceived political difference was a significant predictor of all three measures of relationship quality for student respondents, although the magnitude was smaller than we predicted. In Study 2, when both students and parents were surveyed, we found that perceived political difference did not significantly predict any measure of relationship quality. However, when using simple slopes to examine parent and student responses separately, we observed that for students (though not parents) perceived political difference significantly predicted closeness and was a marginally significant predictor of support. Effect sizes were small, yet these findings build upon those from Study 1 where we found that there is a small but

non-zero relationship between perceived political difference and students' evaluations of relationship quality with their parents.

In both studies, we measured perceived political difference using a single subjective measure, asking if their political views were similar or different. In Study 3, we improved on this by employing two new measures of perceived political difference. Additionally, in Study 3 we build upon open-ended responses that suggest that many students and parents seek to avoid and avert conflict with regards to political conversations by quantifying these patterns and including them in our models.

## Study 3

In Study 3, rather than using a more general measure of perceived political difference, we measure two aspects of perceived political differences separately: *policy views* and *partisanship*.[26] *Policy views* refer to the government policies (e.g., single-payer healthcare, path to citizenship for undocumented immigrants) an individual supports or opposes while *partisanship* refers to the political party an individual identifies themselves with (e.g., Democrat, Republican, or independent). Though related, these aspects of political identity are distinct from each other and have the potential to influence relationships in different ways. Policy differences have the potential to cause disagreements over the implications of policies and what support or opposition to those policies suggests about that person's values. Meanwhile, partisanship can also function as a social identity (Tajfel & Turner, 1979; Greene, 1999; West & Iyengar, 2020) and facilitate prejudice towards out-group members. Although we found little evidence for meaningful relationships between perceived political difference and relationship quality in Studies 1 and 2, we hypothesized that by using these potentially more precise measures in Study

[26] This distinction is similar to what Ellis and Stimson (2012) term "operational ideology" (specific policy views) and "symbolic ideology" (self-identification as liberal or conservative).

3, that these two measures of perceived political difference would predict worse relationship quality. This hypothesis was pre-registered, and we again declared a priori that we would consider this effect to be of a practically meaningful magnitude if it was significantly stronger than a standardized beta of +/− .14 using minimum-effect testing.

Additionally, we added short scales to measure the extent to which students and their parents engaged in conflict avoidance and conflict aversion with regards to political topics. We included these measures as predictors in our pre-registered models, although we did not have an a priori hypothesis about how they would relate to relationship quality. We identified that it is conceivable that when people avoid discussing politics or have to end emotional political conversations that in doing so, they are able to protect their relationships from the harms of conflict. However, it is also possible that avoiding difficult conversations could lead to lower levels of intimacy and openness, and that students and their parents may miss out on important conversations that could strengthen their bond.

## Method

**Participants.** 331 undergraduates (66% women, 34% men, 0.4% nonbinary; $M_{AGE}$ = 19.81, *SD* = 2.97; 44% Asian, 32% White, 8% Latino/a, 8% multiracial, 2% Black, 2% Middle Eastern, 0.8% Native American, 0.4% Native Hawaiian/Pacific Islander) from the University of Washington participated in this study for credit in their psychology courses. They each recruited one of their parents and guardians to participate, and 260 parents/guardians (73% women, 27% men, 0.7% nonbinary; $M_{AGE}$ = 51.41, *SD* = 7.07; 45% Asian, 37% White, 9% Latino/a, 3% Middle Eastern, 2% multiracial, 0.8% Native American, 0.4% Native Hawaiian/Pacific Islander) participated in the study. Only students who indicated they had contact with at least one parent/guardian were eligible to participate. Additionally, in an attempt to have student and

parents who are both familiar with political partisanship in the United States, we only recruited students who did *not* identify as international students. After removing respondents whose parents/children did not complete the survey, we had a total of 250 student-parent/guardian dyads (500 total respondents). Data was collected February-May of 2022.

**Procedure.** The procedure was identical to Study 2 with a few exceptions. First, we added additional questions about participants their political partisanship and their policy views and how they predicted their parents would answer the same questions. Second, participants were also asked questions with regards to political conflict avoidance and aversion. For exploratory purposes, participants were asked if political differences negatively impact their relationship.

**Measures.** Pre-registered models were the same as those in Study 2 with a few exceptions. First, instead of measuring perceived political difference using a single question, we instead asked participants their partisan identities and policy views and asked them to predict their family member's partisanship and policy views. These were then used to calculate each individual's perceived partisan difference and perceived policy difference. Second, we added a dummy coded variable for family role (student or parent). Third, we added measures for conflict avoidance and conflict aversion. Fourth, because a large proportion of respondents in Studies 1 and 2 identified as Asian/Asian-American, we added a dummy coded variable for this identity. New variables are described in detail below.

Additionally, we deviated from our pre-registered analysis plan in two ways.[27] First, we added family role X perceived partisan difference and family role X perceived policy difference to examine if these variables predict student responses differently than parents (as we saw in Study 2). And second, we included moral attributions in our model in order to be consistent with

[27] See Appendix S3 for pre-registered models without deviations from the analysis plan.

the first two studies. However, we did not include interaction terms associated with moral attributions in this study as all moral attribution interactions were non-significant in Study 2, and we did not want to add more predictors to an already large model.

***Perceived Partisan Difference.*** Participants were asked to identify their own political partisanship (Republican, Democrat, or independent) and how they thought their family member would answer that question for themselves. To calculate the level of perceived partisan difference, we compared each individual's party identification to how they thought their family member would identify. If they thought that their family member identified as a different party than them (e.g., a Democrat who thinks their family member would identify as a Republican) this was scored as a 1, an independent who perceives their family member to be partisan or a partisan who views their family member to be an independent was scored as 0.5, and a partisan who thought their family member identified as the same political party was scored as a 0.

***Perceived Policy Difference.*** Participants were asked their opinions on five political issues and how they thought their family members would answer. Those five questions were selected to cover a wide range of political views: single-payer healthcare, government assistance to Black Americans, illegal immigration, abortion, and government services (see Supplemental Materials for full questions). To calculate perceived policy difference, we scaled responses so that the most conservative/Republican response on the scale (e.g., "By law, abortion should never be permitted") was scored as a 1 while the most liberal/Democratic response on the scale (e.g., "By law, abortion should be permitted as a matter of personal choice") was scored as a 0. Answers not on the endpoints of the scale (e.g., "By law, abortion should be permitted only in case of rape, incest, or woman's life in danger") were scored as values between 0 and 1, such that there was equal distance between each value (e.g., a four-point scale would be scored as 0,

.33, .67, and 1). Average perceived policy difference was calculated by averaging the absolute value of the difference between and individual's score and the score for how they perceived their family member would respond for all five policies. Possible values range between 0 (perceived they have the same exact views on all five policies) and 1 (perceive to be on opposite endpoints for all five policies).

***Family Role (Student).*** Dummy coded 1 if respondent is the student, 0 if the respondent is the parent/guardian.

***Political Conflict Avoidance.*** To measure conflict avoidance related to politics, they responded to two questions that were averaged into a two-question scale: "On political issues that I have a different opinion from my [family member], I avoid discussing the topic" (1-7 scale, "strongly disagree"-"strongly agree") and "When my [family member] and I disagree about political issues, we have a conversation about the topic" (same scale, reverse coded).

***Political Conflict Aversion.*** To measure conflict aversion with regards to politics, they responded to two questions that were averaged into a two-question scale: "If my [family member] and I are having a conversation about politics and it is getting emotional, I will try to end the conversation or change the topic" (1-7 scale, "strongly disagree"-"strongly agree") and "If my [family member] and I are having a conversation about politics and it is getting emotional, we continue to have the conversation" (same scale, reverse coded).

***Asian/Asian American.*** Dummy code for if participant identifies as Asian/Asian-American.

**Results and Discussion**

Mean responses for every measure of relationship quality were very similar to measures from Study 2 (see Table 1). The mean perceived policy difference between students and parents

was quite low ($M$ = .17, $SD$ = .15; 0-1 scale) and the majority of respondents perceived they identified as the same political party as their family member (68.9% same party, 22.2% partisan-independent, 8.9% Republican-Democrat). These findings suggest that most respondents perceived themselves to be similar to their family member in terms of policy views and partisanship. Perceived policy difference and perceived partisan difference were moderately correlated with each other, $r$ = .46, $p$ < .001.

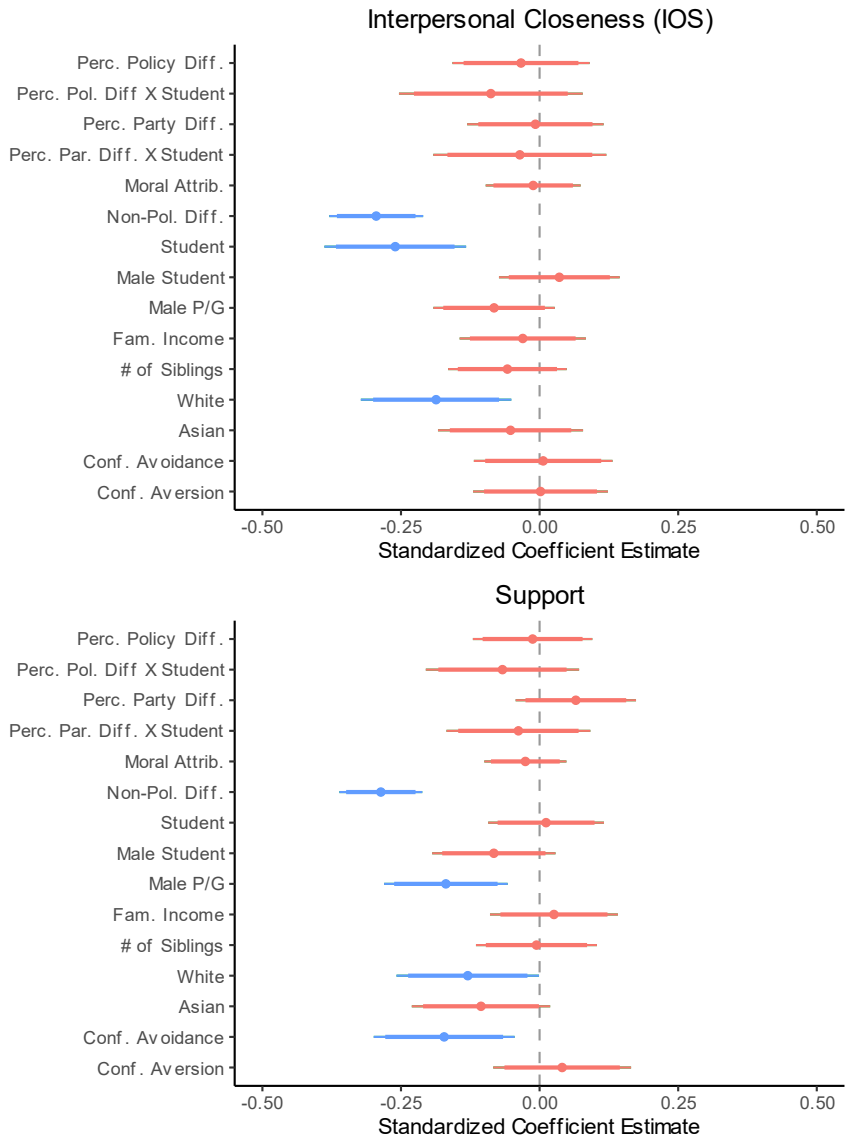
Zero-order correlations were in the hypothesized direction with those who perceived more partisan difference from their family member reporting lower levels of closeness ($r$ = −.13, $p$ = .003) and support ($r$ = −.12, $p$ = .006), and higher levels of negative interactions ($r$ = .12, $p$ = .008). The pattern was similar for zero-order correlations between perceived policy differences and closeness ($r$ = −.16, $p$ < .001), support ($r$ = −.14, $p$ = .002), and negative interactions ($r$ = .17, $p$ < .001).

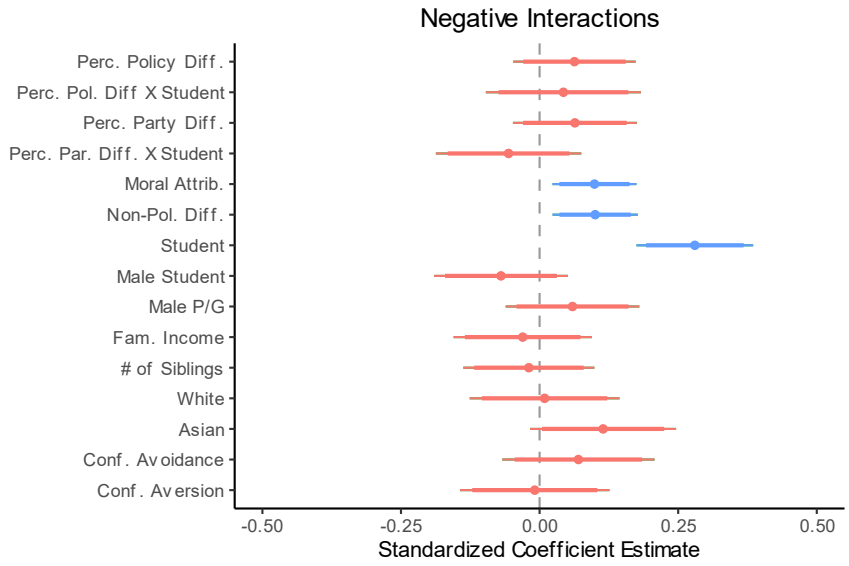
Based on our pre-registration, we used multilevel models with random intercepts for each parent-child pair to test the hypotheses that perceived policy differences and perceived partisan differences are associated with worse relationship quality. We tested three models (see Fig. X), one for each measure of relationship quality. All variables were standardized. Based on a null hypothesis (H0) of no effect whatsoever, we determined that *perceived policy differences* did not significantly predict closeness (β = –.03, 95% CI = [–.15, .09], $p$ = .59), support (β = –.01, 95% CI = [–.12, .10], $p$ = .82), or negative interactions (β = .06, 95% CI = [–.04 .17], $p$ = .26). An equivalence test (pre-registered null region = [–.14, .14]) examining the relationship between perceived policy difference and support suggested that the effect was too small to be practically meaningful, β = –.01, 90% CI = [–.10, .08]. Equivalence tests for the relationship between perceived policy differences and interpersonal closeness (β = −.03, 90% CI = [−.15, .09]) and

negative interactions (β = .06, 90% CI = [−.03, .15]) were not significant, suggesting that there was insufficient data to conclude that the effects are too small to be of practical significance.

**Figure 4**

*Fixed-Effect Predictors of Relationship Quality in Study 3*






*Note*. Standardized β coefficients of fixed effects for three levels of relationship quality. Models had random intercepts for each student-parent dyad. Responses are from both parents and students. Points = estimates. Thick whiskers = 90% confidence intervals. Thin whiskers = 95% confidence intervals. Estimates significantly different than 0 with α = .05 are shown in blue, while non-significant effects are shown in orange. "P/G" refers to parent or guardian.

Using the same models, we also hypothesized that *perceived partisan difference* would predict worse relationship quality (see Fig. 4). We found that perceived partisan differences did not significantly predict closeness (β = –.01, 95% CI = [–.13, .11], $p$ = .59), support (β = .07, 95% CI = [–.04, .17], $p$ = .24), or negative interactions (β = .06, 95% CI = [–.05, .17], $p$ = .43). Equivalence tests (pre-registered null region = [–.14, .14]) examining the relationship between perceived partisan difference and the three measures of relationship quality suggested that the effect was too small to be practically meaningful as a predictor of interpersonal closeness (β =

−.01, 90% CI = [−.11, .09]). Equivalence tests were non-significant for support (β = .07, 90% CI = [–.03, .15]), and negative interactions (β = .06, 90% CI = [–.03, .16]).

Are there differences between students and parents for how perceived political difference and perceived partisan difference predict relationship quality? Family role (student/parent) X perceived policy differences interactions were non-significant for all three measures of relationship quality, suggesting there was not a meaningful difference between students and parents in the predictive power of perceived policy differences. Additionally, family role X perceived partisan differences interactions were also non-significant for all three measures of relationship quality.[28] These findings contrast with those from Study 2, where we observed interactions when predicting closeness and support (but not negative interactions), such that perceived political difference was linked to worse relationship outcomes for student respondents than parents.

It is possible that in Study 3, our measures of partisan difference and policy differences (based on just five policies) did not capture the types of political disagreements parents and students have. There may be differences in approach (practical vs. principled) to solving problems or who to vote for, how important particular issues are perceived to be, or disagreements about how to engage in the political process (protesting, voting, social media posts, etc.). Additionally, disagreements may lie in particular policy issues that were not captured by our measure of policy differences, issues such as gun rights/restrictions, policing, foreign policy, and LGBTQ+ rights.

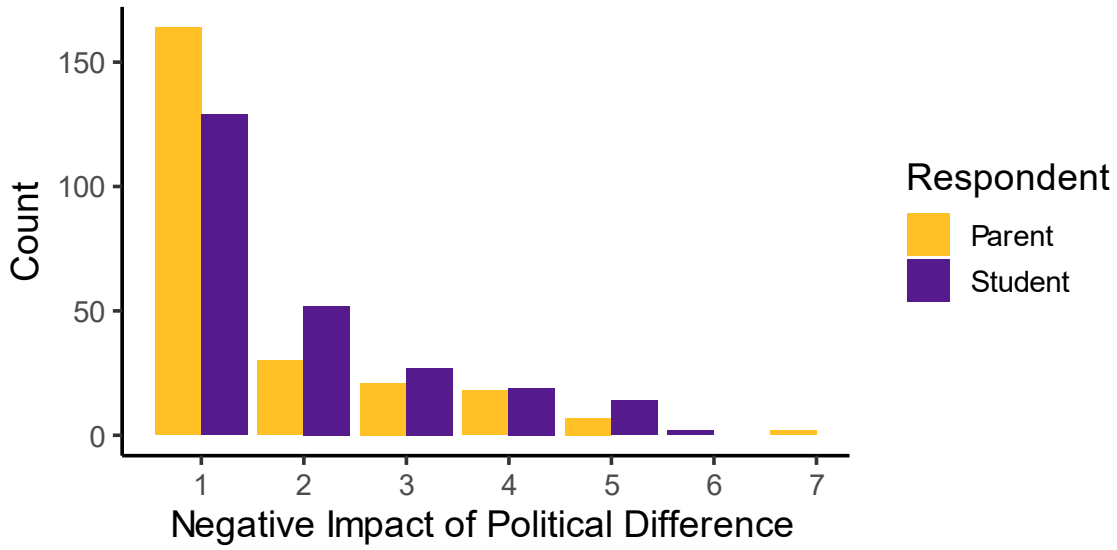
To supplement these correlational analyses that examine how political difference predicts relationship quality, we also directly asked participants how much political differences

[28] Both perceived policy differences and perceived partisan differences were non-significant predictors for all three measures of relationship quality when students and parents were examined separately using simple slopes analyses.

negatively affect their relationship. We found that students (*M* = 1.94, *SD* = 1.27) and parents (*M* = 1.69, *SD* = 1.27) reported generally low levels of negative impact of political differences on their relationship (1-7 scale; not pre-registered). As shown in Fig. 5, the vast majority of respondents reported politics having very little effect on their relationship, with many saying it had no affect at all. However, there were a few respondents with responses near the top of the scale, demonstrating that in some cases politics can be seen as very harmful to parent-child relationships.

**Figure 5**

*Self-Reported Negative Impact of Political Difference on Relationship Quality in Study 3*




*Note.* "To what extent does having different political views from your [family member] affect your relationship negatively affect your relationship with them?" 1 ("not at all") – 7 ("completely") scale.

We did not pre-register a hypothesis about the role of moral attributions in Study 3, but we saw a similar pattern as Studies 1 and 2 for the link between moral attributions for political difference and relationship quality. Like previous studies, there was a main effect of moral attributions predicting negative interactions (β = .10, $p$ = .01, 95% CI = [.02, .17]), which suggests that attributing political differences to moral differences was associated with more negative interactions. As in previous studies, moral attribution was not a statistically significant predictor of interpersonal closeness (β = −.01, $p$ = .79, 95% CI = [−.10, .08]). One point of contrast from Studies 1 and 2 is that we did not observe a statistically significant relationship between moral attributions and support in Study 3 (β = −.03, $p$ = .50, 95% CI = [−.10, .05]).

We explored the role that political conflict avoidance and aversion play in relationship quality (see Figure 4). We identified that they may predict relationship quality, but we did not have an a priori hypothesis about the direction of the effect. On a seven-point scale, average responses were slightly below the midpoint for conflict avoidance ($M$ = 3.38, $SD$ = 1.12) and at the midpoint for conflict aversion ($M$ = 4.00, $SD$ = 1.06). Political conflict avoidance predicted significantly less interpersonal support (β = −.17, 95% CI = [−.30, −.05], $p$ = .008), indicating that those who avoided political conversations offered less support to one another in their relationships. One potential explanation for this finding is that those who do not discuss difficult politics miss out on the opportunity to grow in resolving their differences, but perhaps an even more likely explanation is that those who avoid political conversations do so in part because their relationships are less supportive. However, political conflict avoidance did not significantly predict closeness (β = .01, 95% CI = [−.12, .13], $p$ = .92) or negative interactions (β = .07, 95% CI = [−.06, .20], $p$ = .32). Additionally, political conflict aversion was not a significant predictor of closeness (β = .00, 95% CI = [−.12, .12], $p$ = .98), support (β = .04, 95% CI = [−.08, .16], $p$ =

.56), or negative interactions (β = −.01, 95% CI = [−.14, .12], $p$ = .90). Taken together, these findings cast doubt on the idea that political conflict avoidance and aversion are related to relationship quality in a systematic way.

See Appendix S3 for exploratory analyses examining parent-child transmission rates and perception accuracy by family role and party.

## General Discussion

Based on a growing body of research on affective polarization, we theorized that political differences might be driving a wedge between college students and their parents. We conducted three studies to test whether perceived political difference predicts various aspects of relationship quality for college students and their parents/guardians. We used three different measures of relationship quality in each study and varied our measure of perceived political difference (see Table 2). In all three studies we found that perceived political difference (zero-order) correlated with worse levels of relationship quality. However, when controlling for covariates, this relationship was more complicated. Overall, we found that perceived political difference is associated with worse relationship quality for students in some cases, though the magnitude of the relationship was smaller than expected. Meanwhile, we find no evidence that perceived political differences predict worse relationship quality for parent respondents.

In Study 1 where only students were surveyed, perceived political difference predicted significantly more closeness and support, as well as significantly fewer negative interactions when compared to a null hypothesis of no effect whatsoever. However, when compared to our pre-registered null hypotheses based on the ambient noise level, there was insufficient data to conclude that perceived political difference was related to worse relationship quality to a theoretically meaningful degree. The findings from Study 1 suggest that for college students,

when they perceive that their parents have different political views than them this is associated with worse relationship quality, though the effect was small. Point estimates of effect sizes for the three measures of relationship quality ranged from .11 to .14 standard deviations (standardized beta coefficients). Although we controlled for non-political differences (with a single question) in our models, using the ambient noise level as a baseline allows us to take into account that this study is correlational in nature and that there may be other important factors that covary with non-political difference that we were unable to control for.

In Study 2, we surveyed students and their parents. In this study, perceived political difference (Study 2) did not significantly predict any measure of relationship quality. However, simple slopes analyses revealed that perceived political difference was a significant predictor of lower levels of interpersonal closeness for students, and it was a marginally significant predictor of less support for students as well (though it was not a significant predictor of negative interactions). These findings fall into the pattern observed in Study 1, where even with covariates such as non-political difference, perceived political difference predicted slightly worse relationship quality for students. Unlike student respondents, parents' self-reported relationship quality with their children was not significantly related to perceived political differences.

In Study 3, perceived political difference was measured with two measures, one with regards to perceived *policy* difference and the other regarding perceived *partisan* difference. This change in our primary predictor variable allowed us to examine the effect by using a new measure, and potentially differentiate between differences rooted in identity (partisanship) and views about how government should function (policy). We found that that in our models, neither measure was a significant predictor of any of the three measures of relationship quality. Non-

significant interactions (and simple slopes analyses) suggested that this was the case for both student respondents and their parents.

**Table 2**

*Perceived Political Difference as a Predictor of Relationship Quality in Studies 1-3*

| Outcome Variable | Study 1 | Study 2 | Study 3 | |
|---|---|---|---|---|
| | Perceived political diff. (general) | Perceived political diff. (general) | Perceived *policy* diff. | Perceived *partisan* diff. |
| Closeness (IOS) | β = −.13* [−.23, −.02] | **β = .03 [−.09, .14]** (sig. for students) | β = −.03 [−.15, .09] | **β = −.01 [−.13, .11]** |
| Support | β = −.11* [−.21, −.04] | **β = .03 [−.08, .13]** (marginally sig. for students) | **β = –.01 [−.12, .10]** | β = .07 [−.04, .17] |
| Negative Interactions | β = .14* [.03, .24] | β = .06 [−.04, .17] | β = .06 [−.04, .17] | β = .04 [−.05, .17] |
| Sample | 211 Students (reports on 405 relationships) | 275 Student-Parent Dyads (550 total) | 250 Student-Parent Dyads (500 total) | 250 Student-Parent Dyads (500 total) |

*Note.* *Indicates $p < .05$ using NHST. **Bold** indicates significant equivalence test with null region of −.14 to .14 ("too small to matter"). 95% CIs are in brackets below point estimates. If results varied by family role (parent/student) this was noted in parentheses.

In Studies 2 and 3, where parents also provided responses, we found no evidence suggesting that perceived political difference is related to perceptions of relationship quality with their child. However, according to student responses, perceived political difference is linked to relationship quality in some situations. Yet, even when these effects were significantly different

from zero, they were quite small in magnitude. One potential reason for why perceived political difference and relationship quality are weakly linked (or not linked at all) is that parent responses only came from parents whom, at a minimum, were willing to answer questions to help their child to earn course credit. This selection bias is reflected in the average scores for measures of relationship quality across studies, which show that students in Study 2 and 3 perceived their relationships to be slightly more positive than students in Study 1 where parents were not surveyed. Still, the results from Study 1 where students evaluated their relationships with up to two of their parents (who did not need to complete a survey for their child) still found that the effect was quite small. This casts doubt on the notion that these findings are entirely the result of selection bias.

This study is the first to directly examine the role perceived political differences play in the college student-parent relationship. Given the harmful role of political differences in all types of relationships, it is surprising that the student-parent child relationship does not appear to be meaningfully harmed by perceived political differences. Affective polarization suggests that politics and political identities are a powerful force capable of damaging relationships. Yet, our study provides initial evidence that it may not be the norm for college student-parent relationships. Might the parent-child relationship be one of the few to be able to routinely withstand the effects of affective polarization? The University of Washington, where students were recruited from, is made up of a largely liberal[29] student body which is common among American universities. Still, because we recruited exclusively from one school, future research at a wider range of colleges is needed to be able to generalize to the population of American college students and their parents more broadly. It may be the case that disagreements between conservative family members, or conservative students with more liberal parents, affect the

[29] It is worth noting, however, that there can be a large variety of beliefs even within the same party.

relationship differently than they do in our samples. Additionally, the vast majority of our sample was Asian/Asian-American and White, which further necessitates research on this topic with students of other racial and ethnic backgrounds.

Another finding of our research is that when political differences are attributed to moral differences, this can be associated with worse relationship quality. In Study 1, we hypothesized that there would be a significant interaction such that moral attributions to political differences would be associated with particularly poor relationship quality when there were higher levels of political difference. This hypothesis was supported when predicting negative interactions in Study 1, but not when predicting closeness or support. All moral attributions X perceived political difference interactions were non-significant predictors in Study 2. Although we did not have a hypothesis about the main effect of moral attributions for political differences, we observed that higher levels of moral attributions were associated with less support in the first two studies (though not the third) and more negative interactions in all three studies. These findings are consistent with previous research that shows that when issues are viewed in moral terms that intolerance of divergent views can increase (Skitka, 2010). Future research should seek to probe the ways in which people perceive political views to reflect underlying values, and the ways in which these can influence how people are perceived. Open-ended responses in our research (see Appendix S2) suggest that political views that reflect opinions related to gender identity and sexual orientation may be especially prominent in this regard.

**Conclusion**

This article explored the role that affective polarization plays in the college student-parent relationship. College students and their parents benefit psychologically and practically from maintaining a healthy relationship, and as a result it is important to examine whether

politics can harm this crucial family tie. Despite a growing body of research suggesting that politics often strain and disrupt relationships, we find limited evidence suggesting that the relationship between college students and their parents are meaningfully impacted by perceived political differences. We find that perceiving political differences as reflective of moral differences is predictive of worse relationship quality. Future research is needed before generalizing from these findings, but our research provides a glimmer of hope that the parent-child relationship may be a general exception to the destructive influence of affective polarization.

# ARTICLE 3

## Large enough to matter, too small to be of consequence, or insufficient evidence? A unified framework for a more meaningful statistical inference

Adam H. Smiley, Jessica J. Glazier, and Yuichi Shoda

Department of Psychology, University of Washington, Seattle, WA, 98195

Correspondence concerning this article should be addressed to Yuichi Shoda, University of Washington, 119A Guthrie Hall Box 351525, Seattle, WA 98195-1525.

Email: yshoda@uw.edu

A pre-print of this article can be found at https://psyarxiv.com/465ha/

Keywords: Minimum-Effect Testing; Equivalence Testing; Null Hypothesis Significance Testing; Statistical Inference; Frequentist Statistics

Abstract

Does the observed effect warrant inferring that the true effect in the population is large enough to matter, or too small to matter? Is the evidence too weak to support either inference? The framework presented here answers these questions by considering a “null region,” rather than a point null hypothesis. It also eliminates concerns about “over-powered” studies and facilitates registered reports. This article describes how to: apply this framework without specialized computer programs whenever confidence intervals can be calculated, operationalize “large enough to matter,” and integrate minimum effects testing, equivalence testing, and traditional null hypothesis testing into a single framework.

Traditional null hypothesis significance testing (NHST) deems an observed effect "significant" if it provides strong enough evidence to reject the null hypothesis that there is no effect at all in the population of interest. A major issue with testing against this null hypothesis is that an effect that is *too small to matter* can still be labeled "significant" by traditional NHST. With increased statistical power due to larger samples, NHST will deem even minuscule effects statistically significant. In addition, the null hypothesis is in reference to a *single* value, even though for many purposes, there is a *range* of true effects that are considered not consequential. Furthermore, NHST does not provide a means of "accepting" the null hypothesis with sufficient evidence even when the true effects in the population clearly seem to be in that range. In this paper, we propose a general framework based upon the principles of minimum-effect testing (Murphy & Myors, 1999; Serlin & Lapsley, 1985) and equivalence testing (Dunnett & Gent, 1977; Tryon, 2001; Lakens et al., 2018) to address these shortcomings and allow for more meaningful statistical inference.

In the introduction of this paper, we first provide context to our proposal by describing how others have addressed the issues with traditional null hypothesis testing. Then, we briefly review how minimum effect-testing has been conceptualized in the past, and detail why conducting MET using confidence intervals will make this approach more versatile and accessible. Next, we describe the practice of equivalence testing, how it can work in tandem with MET, and the inferences that can be drawn by using them in a unified framework.

**Null Hypothesis Significance Testing: In Search of an Alternative**

For some research questions, small effects may be of theoretical importance (e.g., Prentice & Miller, 1992; Greenwald et al., 2015). However, in many circumstances it is misleading to refer to very small effects as *significant*, which in common usage means

"sufficiently great or important to be worthy of attention; noteworthy; consequential, influential" (Oxford English Dictionary; Oxford University Press, 2000). Adopting a criterion more stringent than the commonly used $p < .05$ (e.g., $p < .005$; Benjamin, et al., 2018) would increase NHST's evidentiary strength (reflected in the relative odds of alternative vs. null hypotheses, i.e., the Bayes factor), as well as the replicability of statistically significant results. However, even with a different standard, NHST remains fundamentally a test against the null hypothesis of no effect at all. Additionally, the standard practice of conducting null hypothesis testing does not involve *a priori* determination of how effect sizes will be interpreted, leaving room for bias to enter into conclusions drawn based on the size of the observed effect.

With the widely recognized problems associated with NHST such as those described above (see also Greenwald et al., 1996), many have advocated abandoning statistical significance testing altogether in favor of using confidence intervals (e.g., Cohen, 1994). If interpreted properly, confidence intervals are highly informative. In fact, our suggested approach uses confidence intervals as a central part of its decision criterion. However, in practice, researchers often use confidence intervals only to ask if they contain zero, which is functionally no different from traditional null hypothesis testing (Wasserstein & Lazar, 2016). Furthermore, many researchers have "a gross misunderstanding" of the meaning of confidence intervals (Hoekstra et al., 2014).

One reason for the persistent use of NHST may be the heuristic value of a discrete decision criterion. In fact, research on decision making under uncertainty has found that focusing on a discrete consequential outcome while also describing its degree of uncertainty promotes the most effective use of uncertainty information (e.g., Nadav-Greenberg & Joslyn, 2009). Discrete decision criterion is useful for stating the results concisely (e.g., "X had a significant effect on

Y" is more concise than "the effect of X on Y is such-and-such, with a confidence interval of [lower bound, upper bound]"). Authors and editors may also find it useful to have a discrete criterion for deciding what findings are reasonable to be mentioned in the abstract. Is there a way to provide these heuristic values without suffering from the myriad problems of NHST? Below is our attempt to answer this question.

**The Evolution of Minimum-Effect Testing (MET)**

Meehl (1967) argued that when using traditional NHST, increasing precision (e.g., sample size) virtually guarantees that a "statistically significant" result can be found to support any theory even "if the theory is entirely without merit"[30] (p. 111). However, this would not be the case if minimum-effect testing (MET) were used instead. If the observed effect is weaker than the smallest possible effect considered to be meaningful, MET will *never* be significant, regardless of sample size. On the other hand, if the observed effect is stronger than the smallest effect of consequence, the larger the sample size, the greater the likelihood that it is significant by MET.

In addition, minimum-effect testing intrinsically accommodates the fact that often there is a *range* of true effects that are not consequential. This is because MET tests against the null hypothesis that the true effect in the population is within a range of values that are pre-declared as not noteworthy, consequential, or meaningful. It extends the "good-enough principle" advocated by Serlin and Lapsley (1985), in which one first specifies a "belt" around zero and determines if the observed effect falls reliably outside of that range. Murphy and Myors (1999) built upon this idea with minimum-effect tests. In their method of MET, an *F* test is used to

[30] Serlin and Lapsley (1985) took this argument a step further, writing in the abstract, "because the psychological null hypothesis is *always false*, increases in precision in psychology always lead to a weaker test of a theory, whereas the converse is true in physics" (p. 73; emphasis added).

determine if the observed effect is larger than can be reasonably expected from the pre-specified minimally consequential variance accounted for. It will consider an effect significant only when there is sufficient evidence that the observed effect would be unlikely to occur if the true effect in the population was too small to matter. However, in the over twenty years since Murphy and Myors' article was published, MET has been seldom used for empirical research. We suspect this is largely because of the difficulty associated with conducting MET with *F* tests, as well as the lack of versatility offered by the particular method proposed by Murphy and Myors. As we demonstrate below, however, MET can be conducted simply using confidence intervals. This will make MET more versatile, as it can be applied to any statistic for which confidence intervals can be calculated (i.e., not just *F*), and it can be used for both raw and standardized units.

**Equivalence Testing: More than "Retaining the Null"**

If an effect is deemed to be non-significant by MET (or NHST), one cannot simply conclude that the null hypothesis is true. However, by using Equivalence Testing (EqT), researchers can test against the null hypothesis that the true effect in the population is large enough; rejecting this null hypothesis provides a reasonable basis for inferring that the effect is likely too small to be meaningful (e.g., Lakens et al., 2018). Like our proposed method of MET, EqT can also be conducted using confidence intervals (Dunnett & Gent, 1977; Tryon, 2001; Westlake, 1976; Bauer & Kieser, 1996). Much of the literature on Equivalence Testing emphasizes its implementation for determining if two sample means are equivalent (Tryon, 2001) or if an effect is very close to zero (Lakens et al., 2018). However, the logic of equivalence testing can be generalized to a situation where the goal is to provide a basis for inferring that the true strength of the effect is in *any* range, as we discuss later in this article.

**A Unified Framework that Combines MET and EqT**

Both MET and EqT are useful methods on their own, but because they use the same underlying logic, they can naturally be incorporated into an integrated framework. In MET, researchers test if they can infer that the true effect falls *outside* of the range of values that are not of interest. In Equivalence Testing, researchers test if they can infer that the true effect is somewhere *within* the range of values that are not of interest (Lakens et al., 2018). Various terms have been used to refer to this range. In order to use a term applicable to both tests, while emphasizing its similarity to the point null hypothesis in traditional NHST, we refer to it as the "null region." Using MET and EqT in tandem, researchers can draw one of the following three possible inferences about the true effect: (1) the true effect likely falls *outside* of the null region, (2) the true effect likely falls *within* the null region, or (3) the data *do not provide sufficient evidence* to support either inference. This last inference cannot be made by MET alone, or EqT alone. Maxwell and colleagues (2015) have described that data should be looked at from this perspective, and others (e.g., Lakens et al., 2018) have discussed the relationship between MET and EqT. Building upon their recommendations, in this article, we describe how to integrate MET with EqT to optimize the informativeness of statistical inference.

Below, we first conceptualize minimum-effect testing for a directional (one-tailed) hypothesis and then extend it to a non-directional (two-tailed) hypothesis. Second, we describe an overarching framework in which minimum-effect testing is used in conjunction with equivalence testing and explain how this framework can facilitate more meaningful scientific inference. Third, we demonstrate how to conduct MET and EqT using confidence intervals. Fourth, while we believe each scientific community should develop their own standards for specifying the magnitude of effects that would be considered meaningful for a given research question, we offer a few suggestions that could be helpful. Fifth, we discuss how the results of

MET and EqT can be reported succinctly and clearly. And finally, we describe the implications for registered reports and "big data."

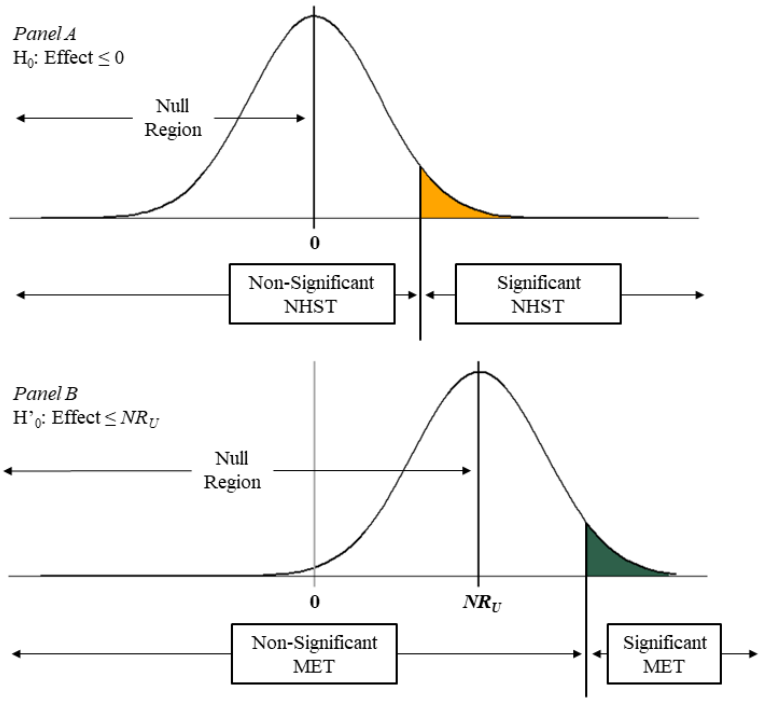
## Conceptualizing One-Tailed Minimum-Effect Testing

The rationale for MET is the same as traditional null hypothesis significance testing (NHST), except for the null hypothesis being tested. Thus, we will first review traditional NHST and then describe how MET differs from it and how to use confidence intervals to perform MET.

The logic of traditional NHST is as follows: (1) We make a counterfactual assumption that, in the population, the true effect is 0 (e.g., the difference between the mean of the experimental condition and that of the control condition is 0). This is the null hypothesis ($H_0$). (2) We then consider the fact that data from any given study are only a subset of the population, and thus the observed value varies from one study to the next and is never *exactly* 0 even if the true effect is 0 in the population. The theoretically expected distribution of the observed values (e.g., sample means) across many repetitions of the study is the sampling distribution shown as the Gaussian "bell" curve shown in panel A of Figure 1. (3) Based on this distribution, we compute the probability of obtaining the observed value or stronger under the null hypothesis. This is the orange area in Panel A of Figure 1. (4) The size of this area, relative to the entire distribution, is the p-value. Conventionally if the p-value is smaller than the value $\alpha$, chosen as the criterion for considering that the observed effect (or stronger) is "unlikely enough" (e.g. $p < .05$, or $p < .005$, as suggested by Benjamin et al., 2018), the null hypothesis is rejected and the observed effect is considered statistically significant.

In one-tailed NHST, the effect is considered "significant" when there is strong enough evidence to reject the null hypothesis that the true effect is 0 or lower (when predicting a positive effect) or that the true effect is 0 or higher (when predicting a negative effect). Thus,

conceptually, the null hypothesis refers to a null *region* of "0 or lower," or "0 or higher", depending on whether a positive or negative effect is predicted. Fortunately, when the observed effect is positive and far enough from 0 to reject the null hypothesis of "absolutely no true effect," it also allows one to reject all other point null hypotheses that the true value is in the null region of "0 or lower," because the p values would have been even smaller if the population value was lower than 0. Similarly, when the observed effect is negative and far enough from 0 to reject the null hypothesis of "absolutely no true effect," it also rejects all other point null hypotheses that the true value is in the null region of "0 or higher."




F*igure 1*. Panel A: A sampling distribution of observed values expected under the null hypothesis $H_0$ that the true effect is zero (traditional NHST). The orange portion represents the area under the curve above the 1 − α quantile (e.g., 0.95) of the sampling distribution. Any observed value in this region is considered unlikely enough given $H_0$, hence statistically significant by NHST. Panel B: A sampling distribution expected under the MET null region hypothesis ($H'_0$) that the true effect is at a value considered the greatest effect of no interest, $NR_U$. The green portion represents the area above the 1 − α quantile for the sampling distribution under $H'_0$. Any observed value in this region is considered significant by MET. Note: In this example the $NR_U$ is shown as larger than the critical value for NHST, but does not necessarily

need to be the case.

The rationale for MET is the same as NHST described above except in a one-tailed MET predicting a positive result, the null hypothesis is that, in the population, there is at most a certain effect pre-declared by the researcher as the largest that is still too small to matter. We will refer to this value as $NR_U$ ("Upper bound of the Null Region").[31] Thus in contrast to the traditional one-tailed NHST for which the null region is "0 or lower" when predicting a positive effect, for a one-tailed MET when predicting a positive effect, the null region is "$NR_U$ or lower." For example, suppose that a public health program is on its way to vaccinate a large enough proportion of the population, such that a new variant of the virus with the reproduction number (the average number of people one infected person will infect) of 3.5 or less would not result in an epidemic, but that a variant with a reproduction number greater than 3.5 would.[32] If a study provides good evidence that the reproduction number of the variant is less than 3.5, then it suggests that the current vaccination program is good enough and no new action is called for. On the other hand, if the study provides evidence that the reproduction number is greater than 3.5 then urgent action is called for. Then for a study assessing the reproduction number of a new variant, it would be reasonable to set the null range for a one-tailed MET as "reproduction number of 3.5 or lower." To distinguish this form of null hypothesis from the traditional null hypothesis ($H_0$) of no effect whatsoever, we refer to it as a *null region hypothesis* ($H'_0$).

Panel B of Figure 1 is the sampling distribution relevant to MET. It is the expected

[31] This example could also be conceptualized as a two-tailed MET (described in the next section) with $NR_U$ = the largest effect considered too small to matter and $NR_L = -\infty$. In bounded distributions, such as correlations (−1 to 1) or percentages (0 to 100%), $NR_L$ would be equivalent to the lowest possible value (e.g., −1 or 0%) rather than $-\infty$.
[32] See, for example, Bartsch et al. (2020). it predicts that if the vaccine has 80% efficacy (i.e., a vaccinated group of people will have 80% less disease cases compared to an unvaccinated group), vaccinating 75% of the population is enough to prevent an epidemic due to a virus with the reproduction number of 3.5 or lower, but that a virus with a reproduction number greater than 3.5 will result in an epidemic even in this relatively well vaccinated population.

distribution of the observed reproduction number across many samples randomly drawn from the population where the true reproduction number is $NR_U$ (e.g., the maximum reproduction number of a virus that the current vaccination program can control). When the green region is "small enough" (i.e., smaller than α), then the observed effect is considered significant using MET. Rejecting $H'_0$ in favor of the prediction of larger effect thus provides evidence that the true effect is large enough to matter. What is "large enough to matter" of course depends on the purpose of research, which should be clearly stated when conducting MET. It is possible that even in the same study, one may report the results of multiple METs with a different null region for each of different purposes, using different values of $NR_U$ as the threshold for "large enough to matter." Multiple METs, each with different null regions, can easily be accommodated in publications (see Table 5 as an example).

## Conceptualizing Two-Tailed Minimum-Effect Testing

Similar to a one-tailed traditional NHST, one-tailed MET only considers an effect "significant" when it is in *one* direction. In contrast, a two-tailed MET considers two directional hypotheses. For example, in a study assessing the reproduction number of a new virus variant, one hypothesis may be that the reproduction number is 3.5 or higher, which prompts urgent action as the current vaccination program will not be able to contain the virus. The other hypothesis is that the reproduction number of the new variant is low enough so that the current vaccination program could be safely halted now before reaching its target population vaccination rate. Suppose this action is warranted if the reproduction number of the new variant is 2.5 or lower.

A reproduction number of between 2.5 and 3.5 does not prompt either action. Thus we may designate the 2.5 to 3.5 range as the null region, in contrast to the "point null hypothesis" of

the traditional NHST. A two-tailed MET tests if a result is significantly higher than the upper bound of the null region ($NR_U$) or if it is lower than the lower bound of the null region ($NR_L$). In this scenario, the virus reproduction number of 3.5 corresponds to the value referred to as $NR_U$ in Figure 2, and the reproduction number of 2.5 corresponds to the value referred to as the $NR_L$ in Figure 2.

In another scenario, for example a study examining the relationship between the level of a newly discovered hormone and mental health, any correlation, positive or negative, with a magnitude of 0.1 may be considered notable. In this case, the null region would span from −0.1 ($NR_L$) to +0.1($NR_U$).

In short, the logic of a two-tailed MET is identical to the one-tailed version except it considers two null hypotheses (corresponding to the two Gaussian curves shown in Figure 3).

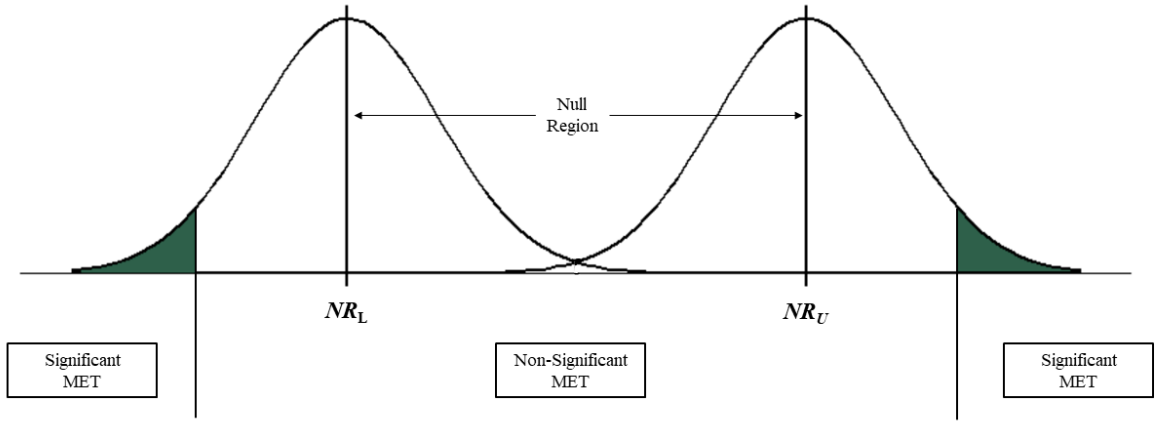



*Figure 3*. Two sampling distributions located at $NR_L$ and $NR_U$, respectively. If an observed effect falls within the shaded region of either distribution (each region = α) it would be considered significant in a two-tailed MET.

Two-tailed MET deems an observed effect significant if it falls in either of the two shaded regions of these curves. That is, two-tailed MET applies the “two one-sided tests”

(TOST) procedure, each with a different directional null hypothesis (Schuirmann, 1987; Lakens et al., 2018). For an alpha level of .05, an effect is deemed significant using a two-tailed MET if it is higher than the 97.5% percentile of the sampling distribution expected from a population at $NR_U$, or if it is lower than the 2.5% percentile of the sampling distribution expected from a population located at $NR_L$.[33]

## Conceptualizing A Unified Framework

When an observed effect is not significant using a two-tailed MET, it may be tempting to conclude that the true effect is located within the null region (i.e., between $NR_L$ and $NR_U$). However, as is the case for any failure to reject null hypotheses, a non-significant effect alone should not be considered sufficient support for *accepting* the null hypothesis.

In this context, it would be helpful to explicitly test if the observed effect is significantly higher than $NR_L$, and *at the same time* significantly lower than $NR_U$. This process is called equivalence testing (Lakens et al., 2018). If they are both significant, this provides good evidence that the true effect in the population of interest is somewhere within the null region.

As shown in Figure 4, this occurs when the following two conditions are *both* met: (1) the observed value is lower than the $\alpha/2$ quantile (e.g., 2.5 percentile) of the sampling distribution for $NR_U$, *and* (2) the observed value is higher than the $1 - (\alpha/2)$ quantile (e.g., 97.5 percentile) of the sampling distribution for $NR_L$.

[33] Strictly speaking, if the true population value is at $NR_U$, there is a non-zero probability of obtaining results that fall in the area on the *left* that is labeled "significant MET." In most situations, the lower critical value is so much farther away from the population value, this probability is virtually 0, making the overall type I error under the null hypothesis nearly entirely in the area under the curve on the right. However, when $NR_L$ is very close to $NR_U$ the sampling distribution from a population located at $NR_U$ may extend meaningfully below the lower critical value. When $NR_L$ and $NR_U$ are identical (i.e., when MET becomes identical to a traditional two-tailed test using NHST), the area under the curve below the lower critical value will equal the area under the curve above the upper critical value. Thus by setting the lower critical value to the $\alpha/2$ quantile of the sampling distribution from $NR_L$ and at the $1 - \alpha/2$ quantile of the sampling distribution from $NR_U$, the type I error rate will be no larger than $\alpha$ (when $NR_L = NR_U$) and no smaller than $\alpha/2$. The same reasoning applies under the null hypothesis that the true population value is at $NR_L$. When the population value is between $NR_L$ and $NR_U$, the type-I error rate is likely much smaller than .05.

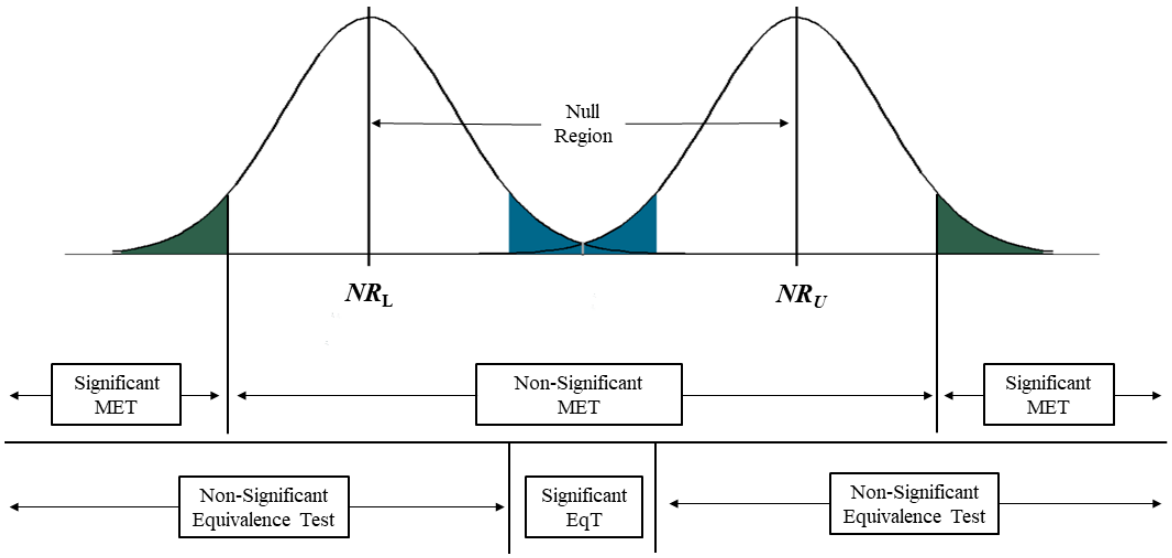



*Figure 4.* MET and Equivalence testing can be used together. If the observed effect is in the blue (central) region, it is significantly lower than $NR_U$ *and* significantly higher than $NR_L$. The green regions (extreme left and extreme right) represent the areas where the effect would be significant by a two-tailed MET, because they are reliably higher than $NR_U$ *or* reliably lower than $NR_L$.

When MET is used together with EqT, there are five potential outcomes corresponding to the five regions in Figure 4, starting from the left most region. This is summarized in Figure 5.

| **Region** | **Region Description** | **Corresponding hypothetical result in Table 1** |
|---|---|---|
| Region 1 | The observed value is significantly lower than the lower bound of the null region (i.e., $NR_L$). | A |
| Region 2 | The observed value is not significant by MET nor by EqT. | B, C |
| Region 3 | The observed value is significant by EqT, offering evidence that it is within the null region. | D |
| Region 4 | The observed value is not significant by MET nor by EqT. | E, F |
| Region 5 | The observed value is significantly higher than the upper bound of the null region (i.e., $NR_U$). | G |

*Figure 5.* Descriptions of and decisions for the five regions shown in Figure 4, starting from the left most region.

When the obtained result is not significant by either test (e.g., it is in region 2 or region 4), researchers should consider the results to be inconclusive and refrain from drawing any conclusions. Note that as sample size increases while the effect size remains the same, Regions 2 and 4, where neither of the tests are significant, will become increasingly narrow.

MET and EqT complement traditional NHST and confidence intervals. For example, results of a hypothetical study assessing the reproduction number of different new virus variants may be reported as follows:

| **Virus variant** | **Observed reproduction number** | **NHST, $H_0$: reproduction number = 0** | **Confidence interval** | **MET and EqT with $NR_L$ = 2.5 and $NR_U$ = 3.5** | | |
|---|---|---|---|---|---|---|
| | | | | MET: reproduction number ≤ 2.5 | EqT: 2.5 < reproduction number < 3.5 | MET: reproduction number ≥ 3.5 |
| **A** | 2.0 | p<.001 | 1.7-2.3 | p<.001 | ns | ns |
| **B** | 2.3 | p<.0005 | 2.0-2.6 | ns | ns | ns |
| **C** | 2.7 | p<.0005 | 2.4-3.0 | ns | ns | ns |
| **D** | 3.0 | p<.0001 | 2.7-3.3 | ns | p<.001 | ns |
| **E** | 3.3 | p<.00005 | 3.0-3.6 | ns | ns | ns |
| **F** | 3.7 | p<.00001 | 3.4-4.0 | ns | ns | ns |
| **G** | 4.0 | p<.000001 | 3.7-4.3 | ns | ns | p<.001 |

*Table 1*. Hypothetical results of a vaccine efficacy study. ns = not significant.

### Conducting Minimum-Effect and Equivalence Testing with Confidence Intervals

MET and EqT can be conducted if the statistic of interest has a method for determining confidence intervals, even without explicitly needing to determine the sampling distribution (this is because the computation of confidence intervals already takes sampling distributions into account, see Appendix S1: *The Rationale for using Confidence Intervals to conduct MET* for an explanation). The procedure is described in Table 2.

How to Conduct Minimum-Effect Testing (MET) and Equivalence Testing (EqT) using confidence intervals

---

One-tailed MET, when *predicting that the effect is* **higher than the null region**, is significant when:

- the confidence interval is entirely **above** $NR_U$, the highest population value that is still too low to be of consequence.

---

One-tailed MET when *predicting that the effect is* **lower than the null region**, is significant when:

- the confidence interval is entirely **below** $NR_L$, the lowest population value that is still too high to be of consequence.

---

Two-tailed MET is significant when the result is **significantly outside of the null region**. i.e.,

- the confidence interval is entirely **above** $NR_U$, the highest population value that is still too low to be of consequence.

  **or**

- the confidence interval is entirely **below** $NR_L$, the lowest population value that is still too high to be of consequence.

---

Equivalence testing is significant when the result is **significantly within the null region**. i.e.,

- the confidence interval is entirely **below** $NR_U$, the highest population value that is still too low to be of consequence.

  **and**

- the confidence interval is entirely **above** $NR_L$, the lowest population value that is still too high to be of consequence.

*Table 2.* A summary of how to conduct a one-tailed or two-tailed MET and an EqT using confidence intervals. $NR_L$ and $NR_U$ refer to the value(s) of lower and upper bounds of the null region.

## Specifying the Null Region

Specifying $NR_L$ and $NR_U$ plays a critical role for MET and EqT. As it should be guided by specific research questions and involves considering many factors, an important challenge for any field of study (or scientific journal) is developing practices and standards for specifying them. For example, clinical psychologists may be most interested in considering clinical significance, but public health researchers may find that utility analysis often makes the most sense. Below, as a guide, we summarize a number of approaches that could be used to specify meaningful $NR_L$ and $NR_U$ values.

Figure 6 presents possible bases for $NR_L$ and $NR_U$ values and describes the rationale for each. They are not meant to be exhaustive; there may be other means of specifying $NR_L$ and $NR_U$ values that have not yet been suggested. We encourage each scientific discipline to develop their own guidelines and practices most applicable to their research question.

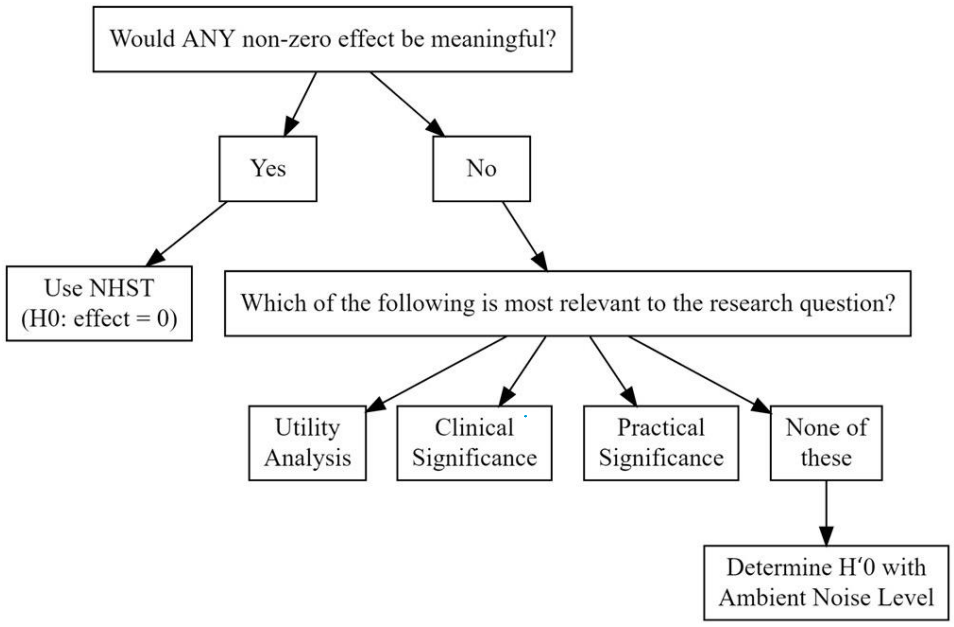



*Figure 6.* This flowchart provides a series of questions researchers could ask when determining the $NR_L$ and $NR_U$.

**Any non-zero true effect is meaningful.** In some circumstances, *any* non-zero true effect might be meaningful, in which case $NR_L$ or $NR_U$ would be 0 and MET becomes equivalent to NHST. Researchers, however, should explicitly state and justify why any true non-zero effect would be meaningful.

**Utility Analysis.** When Murphy and Myors (1999) described their procedure for minimum-effect testing, they suggested the use of utility analysis in determining a minimum effect. Utility analysis involves comparing the costs of an intervention against their benefits measured in "utility based units" (Robinson, 1993). This utility is often described in terms of financial consequences (Schmidt et al.,1979), or in terms of patient outcomes such as quality adjusted life years (Robinson, 1993). Regardless of the units used, if $NR_L$ or $NR_U$ is set at the magnitude of the effect where the cost and benefit of the effect "break even," a significant MET suggests that the benefit outweighs the costs.

**Clinical Significance.** In the context of clinical intervention research, traditional statistical significance testing has long been criticized for its failure to accurately reflect clinically meaningful change, because a statistically significant finding is not necessarily clinically meaningful. Jacobson and colleagues (1984) first proposed clinical significance in the context of clinical practice as a measure of whether a client has moved from the dysfunctional to functional range on the assessment being used to track change during treatment. When considering clinical significance in determining $NR_L$ and $NR_U$, we suggest that researchers take into account the Minimal Important Difference (MID; see King, 2011 for a discussion of related terms and methods of setting MID).

**Practical Significance.** Clinical significance can be thought of as an evolution of a more

general concept of practical significance (Pintea, 2010). For the purpose of conducting MET, $NR_L$ and $NR_U$ can be the smallest true effect size that would be of potential utility in *practice* or in the real world (Kirk, 1996).

**Ambient Noise Level.** For many research questions, for example in research where there is no obvious cost to weigh against outcomes, the approaches above may not be feasible or appropriate in determining $NR_L$ and $NR_U$. Under such circumstances, one reasonable basis for setting $NR_L$ and $NR_U$ is the "ambient noise level" (Lykken, 1968; also called also called the "crud estimate," Orben & Lakens, 2019). This refers to the fact that even in extremely large samples, non-zero correlations exist between variables with no obvious relation to each other (e.g., interest in woodworking and birth order). This led Paul Meehl to observe that "In the social sciences and arguably in the biological sciences, 'everything correlates to some extent with everything else'" (Meehl, 1990, p. 204). That is, because of associations due to historical accidents (e.g., Krugman, 1991) rather than to any intrinsic or theoretically meaningful relationship, the correlation between two variables that are conceptually not related at all will still approach a significant value with an ever-increasing sample size. Similarly, in an experiment, even if the theoretical variable intended to be manipulated has no true effect at all, with an ever-increasing sample size the effects of any actual experimental manipulation will asymptote towards a non-zero value, because it is virtually impossible to manipulate only the theoretical variable of interest (i.e., it is impossible to eliminate all confounds).

Several scholars have attempted to estimate the size of the ambient noise level by looking at large correlation matrices. Pearson's $r$ of .07 (corresponding to Cohen's $d$ of 0.14) seems to provide a reasonable approximation for correlations between variables for which there is no theoretical reason to expect any correlation. This could be used as a baseline across disciplines

(see Standing et al.,1991, who based this estimate on correlations computed with a sample size over 2000; see also Webster & Starbuck, 1988). Thus, when it is not possible to determine $NR_L$ and $NR_U$ via other methods, $NR_L$ and $NR_U$ could be set at least as large as the ambient noise level. For example, the null region could be set between $r = -.07$ and $r = +.07$ (or $d \leq 0.14$). Doing so can reduce the likelihood of labelling relationships reflecting historical accidents as theoretically meaningful.

### *A priori* determination of the Null Region

We suggest that $NR_L$ and $NR_U$ values be determined and pre-registered prior to data collection/analysis whenever possible. Without determining $NR_L$ and $NR_U$ prior to data analysis, researchers may be more likely to fall into committing *post hoc* rationalizations of small effect sizes, or *HARKing (Hypothesizing After the Results are Known*, Kerr, 1998). Researchers can utilize pre-registration as a way to show that their $NR_L$ and $NR_U$ values were determined *a priori*.

### Considerations for Editors and Reviewers

Although we hope that researchers will be unbiased in determining $NR_L$ and $NR_U$, it is important to consider the incentives that affect researchers. Many researchers may be motivated to set $NR_L$ and $NR_U$ values to specify an excessively narrow null region, out of concern that setting a wide null region would make it difficult to reject the MET null hypothesis. For this reason, editors and reviewers should consider the appropriateness of the $NR_L$ and $NR_U$ values chosen by researchers. We believe there are a number of potential mechanisms that can be used to address this issue (for an example, see Appendix S2: *Peer-Review*), such that the balance of the authors' and reviewers' incentives, in addition to their commitment to the fundamental goals of science, will result in constraining the proposed $NR_L$ and $NR_U$ values within a reasonable range. This process will also facilitate a benefit of using MET, namely increasing discourse

about effect sizes as well as transparency in the research process.

## Reporting Findings

We have provided three examples of how MET may be used in reporting findings. These examples demonstrate how this can be done concisely, and how it can be done in addition to, rather than instead of, the results of traditional NHST.

### Example 1

Table 3 shows an example of the (hypothetical) results from a two-group experiment using one-tailed MET and EqT. Specifically, this study tested the hypothesis that when Democrats and Republicans undergo a "fast friends" paradigm to get to know each other (Aron et al., 1997) their level of warmth towards members of the other political party (the "DV" in Table 3) will increase meaningfully (effect size ≥ .2) in comparison to a control intervention.

| Experimental Condition | DV | Effect Size (EF) | Confidence interval | NHST | MET and EqT | |
|---|---|---|---|---|---|---|
| | | | | | EqT<br>−0.2 < true EF < +0.2 | One tailed MET<br>0.2 ≤ true EF |
| **Fast Friends (FF)** | **3.5** | | [3.1, 3.9] | | | |
| **Control (C)** | **2.4** | | [2.1, 2.7] | | | |
| **Difference** | | **.7** | [0.3, 1.1] | p<.001 | ns | $p < .01$ |

*Table 3*. The hypothetical results of a two-group experiment analyzed with MET and EqT.

### Example 2

Table 4 provides an example of a linear regression from a simulated "big data" study and how results could be interpreted using a two-tailed MET and EqT. N = 10,000 and H'$_0$s ($NR_L$ and $NR_U$) vary by predictor variable (see H'$_0$ column).

| Predictor | Standardized beta | SE | t-value | p-value (NHST) | 95% CI | H'$_0$ | MET/EqT |
|---|---|---|---|---|---|---|---|

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Intercept** | 0.000 | .009 | 0 | 1 | [−.017, .017] | N/A | N/A |
| **A** | 0.300 | .009 | 34.44 | < .001 | [.283, .317] | −.1, .1 | Sig. MET |
| **B** | 0.009 | .009 | .21 | .21 | [−.006, .027] | −.1, .1 | Sig. EqT |
| **C** | 0.390 | .009 | 44.79 | <.001 | [.373, .407] | −.2, .2 | Sig. MET |
| **D** | 0.108 | .009 | 9.18 | <.001 | [.091, .125] | −.1, .1 | n.s. |

*Table 4.* The results of linear regression with N = 10,000 and varying H'$_0$s by predictor variable.

**Example 3**

Below, the description of the results reported in Table 4 of Shoda et al. (1990) has been modified to include results of NHST, MET, and EqT simultaneously.

> For MET and EqT, two sets of analyses were conducted, corresponding to two possible purposes for which the Shoda et al. (1990) findings are likely to be used. The first purpose is testing the theoretical proposition that individual differences in waiting time in the diagnostic condition reflect something relatively lasting (e.g., cognitive skills, family, or community environment conducive to school achievement). For this purpose, any evidence that the correlations between waiting time and future outcomes are more than what are expected for any randomly chosen pair of variables would be meaningful. Thus, we used the commonly accepted "ambient noise level" of $r = -.07$ and $r = .07$ as the $NR_L$ and $NR_U$ respectively. Against this null hypothesis, the two correlations that were significant by NHST (the ones with asterisks in Table 4) were also significant with a two-tailed MET.
>
> The second purpose for which the present results can be used involves practical consequences of the differences in SAT scores predicted from preschool waiting time. Thus, in this second set of MET and EqT, we based the H'$_0$ values on practical significance, in particular the expert opinion published in *New York Times* that a difference of as few as 20 points on a student's SAT could make a difference in college admissions or financial aid (Aviv, 2009). Based on this, we chose a 20-point difference in SAT as the smallest effect of interest (positive or negative). How strong do the correlations need to be in order to predict a 20-point difference in SAT scores if one child waits 5 minutes longer than another? We used the additional information from the same data set reported on page 936 of Mischel, Shoda, & Rodriguez (1989) and determined that if correlations between preschool waiting time and SAT are at least 0.139 for SATV, and 0.145 for SATQ, then a difference of 20 points on SATV and SATQ, respectively, is predicted for every 5 additional minutes of waiting time in preschool. Thus, in this set of MET and EqT, for the observed correlations of SATV and SATQ with preschool waiting time, correlations of +/− 0.139 (for SATV) and +/− 0.145 (for SATQ) were used as $NR_L$ and $NR_U$ values.

Below (Table 5 in this manuscript) is how Table 4 from the original Shoda et al. (1990) article would be revised, in light of the above. No deletions from the originally published table were made; additions related to MET and EqT are underlined.

Correlations Between Preschool Delay Time and Scholastic Aptitude Test (SAT) Scores

| Measure | Spontaneous ideation | | Suggested ideation | |
|---|---|---|---|---|
| | Rewards exposed | Rewards obscured | Rewards exposed | Rewards obscured |
| SAT Verbal | .42* | −.12 | −.40 | −.21 |
| SAT Quant. | **.57**** | −.31 | −.26 | −.23 |
| Sample size | 35 | 33 | 14 | 12 |

*Table 5.* *: $p < .05$, and **: $p < .01$ by NHST with $H_0 : r = 0$. **Bold:** $p < .05$ by two-tailed MET with $NR_L$: $r = -.139$, $NR_U$: $r = +.139$ for SATV and $NR_L$: $r = -.145$, $NR_U$: $r = +.145$ for SATQ, based on considerations of practical significance. Another set of MET, conducted against the null hypothesis that the effect is no greater than a common estimate of "ambient noise" indicated that the two correlations that were significant by NHST were also significant with a two-tailed MET. None of the correlations were close enough to 0 to be significant by Equivalence Testing using the same $H'_0$s.

The following is how the findings may be referred to in the discussion section, in light of MET and EqT. The passage is taken verbatim from page 985 of Shoda et al. (1990). No deletions were made; additions related to MET and EqT are underlined:

> In addition, given the smallness of the sample, the obtained coefficients could very well exaggerate the magnitude of the true association. For example, in the diagnostic condition, the 95% confidence interval for the correlation of preschool delay time with SAT verbal score ranges from .10 to .66, and with SAT quantitative score, the confidence interval ranges from .29 to .76. Thus, using a two-tailed MET, only the correlation with SAT quantitative is significant against the null hypothesis that the true magnitude of the correlation is large enough to predict a meaningful practical difference for a college applicant (at least a 20-point difference in SAT score for every additional 5 minutes in preschool waiting time). Although the magnitude of the observed correlation with SAT verbal was large enough to reject the traditional null hypothesis that the true correlation is 0 as well as the null hypothesis that the true correlation is at or weaker than "ambient noise," it was not large enough to reject the null hypothesis that the true correlation is large enough to make a practical difference for college applicants. In the other conditions,

> none of the correlations were significant with MET or EqT against either null hypothesis. Thus, overall, except for the correlation with SAT quantitative in the diagnostic condition, the correlations were not large enough to make a practical difference. Nor was it small enough to reject the null hypothesis that the true correlation is no larger than the smallest effect that still matters. In addition**,** [t]he value and importance given to SAT scores in our culture make caution essential before generalizing from the present study; at the very least, further replications with other populations, cohorts, and testing conditions seem necessary next steps…

Below is the entirety of the Shoda et al. (1990) abstract, edited in light of MET (deletion shown in strikeout, and additions are underlined):

> Variations of the self-imposed delay-of-gratification situation in preschool were compared to determine when individual differences in this situation may predict aspects of cognitive and self-regulatory competence and coping in adolescence. Preschool children from a university community participated in experiments that varied features of the self-imposed delay situation. Experimental analyses of the cognitive-attentional processes that affect waiting in this situation helped identify conditions in which delay behavior would be most likely to reflect relevant cognitive and attentional competencies. As hypothesized, in those conditions, coherent patterns of statistically significant correlations were found between seconds of delay time in such conditions in preschool and cognitive and academic competence and ability to cope with frustration and stress in adolescence. In the "diagnostic" condition, the correlation between participants' preschool delay time and SAT quantitative score was significant against the pre-registered null hypothesis that the true correlation is r = .145 (or smaller), which is the minimum that may predict a difference of 20 SAT points for every 5 additional minutes of preschool waiting time; the correlation with the verbal score was not strong enough to suggest such practical difference, although it was still significant against the traditional null hypothesis that the true correlation is 0. In the other conditions, the results did not provide sufficient evidence for any conclusion of potential interest; we cannot conclude they are large enough to make a practical difference, nor can we conclude that they are practically equivalent to 0.

As shown by these examples, reporting MET can concisely add important information about the magnitude of observed effects without sacrificing what is already conveyed using traditional NHST.

## Registered Reports

MET solves an important challenge for the promising new publication mechanism called the *registered report*. Registered reports are a publishing format which conduct peer review prior

to data collection. Publication of the results is provisionally accepted as long as the authors closely adhere to the registered methodology (Center for Open Science, n.d.). It minimizes a variety of questionable research practices, such as publication bias and selective reporting of "significant" or hypothesis-confirming results, while making it possible for reviewers to "help authors to improve the protocol or rationale while it is still possible to make changes" (Chambers, 2019). One important question, however, has remained: How does one define that a prediction has been confirmed? Simply stating that predictions are confirmed when there is a statistically significant effect using NHST is problematic because as long as the true effect is not exactly zero, traditional NHST statistical significance is all but guaranteed as the sample size increases (in a two-tailed test). Considering that the prediction is confirmed when the observed effect was larger than the pre-registered minimum predicted effect is also problematic. This is because the likelihood of such an outcome is higher, rather than lower, with a smaller sample size (hence greater standard error), in effect incentivizing researchers to opt for smaller sample sizes.

The proposed framework integrating MET and EqT avoids both of these problems. For example, suppose in a registered report authors predicted that a new housing policy will reduce the number of people experiencing homelessness in the population by at least 30%. Unlike NHST, unless the reduction observed in a study is greater than 30%, the results will never be statistically significant by MET no matter how large the sample is. Smaller sample size and the resultant larger standard error will increase the probability of observing reductions greater than 30% by chance, but because MET takes the larger standard error into account, the false positive rate (i.e., Type I error) will be capped at the chosen level of $\alpha$ (e.g., 0.05). For an example of MET in a registered report, see King et al. (in press).

## Implications for Sample Size

Open science practices increasingly recognize the importance of adequately large sample size. At the same time, it has been noted that when using traditional NHST, increasing the sample size will make it more likely to reject the NHST null hypothesis as long as the true effect is not exactly 0. Results of studies using exceptionally large sample sizes can become unintentionally misleading if one only considers whether the NHST p-value is statistically significant. Unlike traditional NHST, however, when using the framework described above, larger sample sizes do not result in misleading conclusions. For example, if the observed effect falls within the null region, it will *never* be significant by MET regardless of sample size (to see why this is the case, note that as the sample size increases and the "width" of sampling distributions decreases in Figure 1, any observed effect between 0 and $NR_U$ will eventually be in the area shown in orange, hence significant by a traditional NHST. However, the same observed values will never be in the green area (values that are significant by MET), no matter how narrow the sampling distributions become). On the other hand, if the observed effect falls outside of the null region, the larger the sample size is, the more likely that it is significant by MET. Similarly, increasing sample size will increase the likelihood of significant EqT only if the effect is within the null region. Another way to describe the consequence of increasing sample size is that it would decrease the region in which the results are inconclusive, allowing researchers to accurately conclude either the effect to be strong enough for a significant MET, or weak enough to be equivalent to 0 (or in the null region, more generally).

## Conclusion

With the present framework, researchers can infer, with greater confidence, if an effect is large enough to matter, too small to be of consequence, or if neither inference is warranted. It

avoids many well-documented problems stemming from always testing against the null hypothesis of "no effect," even in a research context in which rejecting such a null hypothesis is akin to a proverbial attack on a "straw man." And it provides a language for stating the inference very concisely (e.g., "significant by MET"), for example in abstracts where brevity is paramount, while requiring researchers to explicitly spell out the basis for considering an effect as meaningful (i.e., confidence intervals and the range(s) of population values pre-declared as meaningful, reported in the main body of the paper). The proposed framework does not result in labeling small but non-zero effects as "significant" if they are too small to matter, even in research involving "big data." Additionally, when an effect is very small, it allows researchers to test if the true effect is in fact too small to matter. This can facilitate the publication of "null" results, mitigating the "file drawer" problem in the publication process (a bias against publishing findings that are not statistically significant). It supports the emerging practice of registered reports by providing a much-needed method for stating, and testing, predictions about the true magnitude of an effect. This avoids incentivizing smaller sample size, as would be the case if the criterion is stated in observed (rather than population) values. In sum, we believe that the framework proposed here can lead to improvements in the quality of data-based inferences in scientific research.

## Conclusions and Future Directions

In this dissertation, I examined the causes and implications of affective polarization. We identified that geographic sorting contributes to affective polarization in the United States. We suggest that people who live in politically homogenous communities aligned with their political identities have higher levels of affective polarization due to a lack of cross-party contact. Furthermore, we demonstrate that cross-party contact can reduce affective polarization for people in political bubbles.

Amidst an extensive body of literature which suggests that politics can drive a wedge in relationships, we examined the role perceived political difference plays in relationships between college students and their parents. Contrary to our expectations, we found that the association between perceived political difference and relationship quality are minimal, when controlling for other factors such as non-political differences. There does not appear to be any link between perceived political differences and relationship quality for parents, but student responses suggest a small but non-zero association between perceived political difference and relationship quality. Our investigation was aided by minimum effect testing, the subject of the third article of this dissertation, which can be a powerful tool for improving statistical inference and encouraging better research practices.

Overall, my research suggests that cross-party contact is important for reducing affective polarization, and that at least within the parent-child relationship it is unlikely that political differences will cause meaningful relational harm. The parent-child relationship may be exceptional in this regard, but it provides a glimmer of hope that relationships across the political divide can remain intact and provide a vessel for the ameliorating effects of cross-party contact. This research points to several lines of future investigations, which I detail below.

## Future Directions

### Love the Sinner, Hate their Group?

One of the surprising findings from our research on geographic sorting and affective polarization is the role of church attendance on affective polarization. In Study 1, where we used a multiple regression model to predict *party*-directed affective polarization, church attendance was not a significant predictor. However, in Study 2, more frequent church attendance predicted significantly less *voter*-directed affective polarization and significantly less social distance. Party-directed affective polarization is generally higher than voter-directed, as it reflects attitudes towards political elites in addition to average party supporters (Druckman & Levendusky, 2019). In Study 3, republican participants who were students at Christian universities reported much less voter-directed affective polarization ($M$ = 13.59) than party-directed ($M$ = 30.36), while democratic participants from a public university didn't show this difference (Voter-Directed $M$ = 46.63, Party-Directed $M$ = 47.50).

Taken together, these findings suggest that there may be something about church attendance or Christianity more broadly that may lead to lower levels of voter-directed affective polarization and social distance, though not necessarily party-directed affective polarization. This appears to reflect Jesus' call to love one's enemies (Matthew 5:43-48) or the adage to "hate the sin, love the sinner."[34] Attending church may also present an opportunity for cross-party contact, although it is unclear why this form of contact would differentially affect voter- and party-directed affective polarization to this extent. Future research should be conducted to examine the role religion and church attendance play in shaping attitudes towards outpartisans.

### Measuring Affective Polarization and Predicting Behavior

[34] This assumes that some people think of supporting a certain political party as sinful… which feels like a safe bet.

For decades, affective polarization has been measured using party-directed feeling thermometers (e.g., Iyengar et al., 2019; Mason, 2018). Social distance has also been measured similarly over the years, with a particular emphasis on the question of how comfortable people would be having their child marry someone from the out-party. These consistent measures have allowed for researchers to examine trends over time, and for that reason it makes sense to continue using them. However, questions remain about the quality of these measures and the applicability for predicting behavior.

Affective polarization as measured by feeling thermometers is calculated by taking the difference between in-party warmth and out-party warmth. Are evaluations of the in-party relevant for how cross-party interactions transpire? Or are only perceptions of the out-party relevant? Additionally, as we have seen, asking participants to indicate how warm they are for parties vs. voters can lead to quite different outcomes in some cases (see also Druckman & Levendusky, 2019). It is likely that which measure is most useful depends on the research question, but more research is needed to understand how these measures relate to actual behavioral outcomes. In the future, I hope to examine how these measures of affective polarization relate to how Democrats and Republicans interact with one another, as well as working to develop new measures of affective polarization.

**LGBTQ+ Perceptions of Political Views**

In our investigation of how perceived political differences relate to parent-child relationship quality, we found that political differences were rarely associated with worse relationships. However, in response to open-ended questions, a number of LGBTQ+ students indicated that they were scared to come out to their parent because of their parents' political views. Others described that they tried to avoid discussing politics because they feared

accidentally outing themselves, or because they wanted to avoid having their feelings hurt by hearing their parents' opinions about LGBTQ+ issues. Parental support is particularly important for the mental health of LGBTQ+ adolescents (e.g., Mills-Koonce et al., 2018; Ryan et al., 2009, 2010), so these differences in political views which may signal a lack of support for a stigmatized identity may cause real harm. Future research should be conducted to examine if political differences are particularly harmful to relationships for LGBTQ+ individuals.

**Are Political Beliefs Perceived as Reflections of Moral Values?**

Relatedly, we found that when people attribute differences in political views to differences in morality, relationship quality was worse. Moral convictions can lead to stronger emotional responses when considering a topic (Skitka et al., 2005) and intolerance of those who hold different views increases when a topic is viewed as moral in nature (Skitka, 2010). Future research should consider the extent to which people perceive the political views of others to reflect that person's moral values (e.g., "Steve thinks we should build a wall at the southern border, he must be a racist") or whether they perceive a person's views as instrumental in leading to policy change (e.g., "Steve thinks we should build a wall at the southern border, Steve and others like him may cause a wall to be built, which I wouldn't like"). It is likely a combination of both, but to varying degrees. Opinions about issues which are especially emotionally charged (e.g., abortion, LGBTQ+ rights, immigration, gun control, etc.) may be seen as more reflective of a person's moral values than other issues (e.g., tax rates, tariffs, etc.). Future research is needed to examine the specific mechanisms by which knowledge of a person's political views shapes how they are perceived.

# References

Abrams, S. A. (2021, Sep. 20). Are American families really being torn apart by politics? *Institute for Family Studies.* https://ifstudies.org/blog/are-american-families-really-being-torn-apart-by-politics

Allport, G. W. (1954). *The nature of prejudice.* Perseus Books.

Ambert, A. M. (1997). *Parents, children, and adolescents: Interactive relationships and development in context*. Psychology Press.

American National Election Studies. (2017). ANES 2016 Time Series Study (ICPSR 36824). Ann Arbor, MI: Inter-university Consortium for Political and Social Research. doi:10.3886/ICPSR36824.v2

American National Election Studies. (2021). ANES 2016 Time Series Study Geocodes (ICPSR 38087). Ann Arbor, MI: Inter-university Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR38087.v1

Amsalem, E., Merkley, E., & Loewen, P. J. (2022). Does Talking to the Other Side Reduce Inter-party Hostility? Evidence from Three Studies. *Political Communication*, *39*(1), 61-78.

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and social psychology bulletin*, *23*(4), 363-377.

Aviv, R. (2009, October 26). Proctor, Please Get Off the Phone. *New York Times*.

Barnett, R. C., Marshall, N. L., & Pleck, J. H. (1992). Adult son-parent relationships and their associations with sons' psychological distress. *Journal of Family Issues*, *13*(4), 505-525.

Bartsch, O'Shea, K. J., Ferguson, M. C., Bottazzi, M. E., Wedlock, P. T., Strych, U., McKinnell, J. A., Siegmund, S. S., Cox, S. N., Hotez, P. J., & Lee, B. Y. (2020). Vaccine Efficacy

Needed for a COVID-19 Coronavirus Vaccine to Prevent or Stop an Epidemic as the Sole Intervention. *American Journal of Preventive Medicine, 59*(4), 493–503. https://doi.org/10.1016/j.amepre.2020.06.011 .

Bauer, P., & Kieser, M. (1996). A unifying approach for confidence intervals and testing of equivalence and difference. *Biometrika, 83*(4), 934-937.

Benjamin, Daniel J, Berger, James O, Johannesson, Magnus, Nosek, Brian A, Wagenmakers, E-J, Berk, Richard, . . . Johnson, Valen E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6-10.

Bishop, B. (2020, December 17). For Most Americans, the Local Presidential Vote Was a Landslide. *Daily Yonder*. Retrieved March 11, 2022, from https://dailyyonder.com/for-most-americans-the-local-presidential-vote-was-a-landslide/2020/12/17/.

Bishop, B., & Cushing, R. G. (2008). *The big sort: Why the clustering of like-minded America is tearing us apart.* Mariner Books.

Bogardus, E. S. (1933). A social distance scale. *Sociology & Social Research*.

Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. Basic books.

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2020). *Cross-country trends in affective polarization* (No. w26669). National Bureau of Economic Research.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate?. *Journal of Social Issues*, *55*, 429-444.

Brown, J. R., & Enos, R. D. (2021). The measurement of partisan sorting for 180 million voters. *Nature Human Behaviour*, *5*(8), 998-1008.

Center for Open Science (n.d.). Registered Reports: Peer review before results are known to

align scientific values and practices.

Chambers, C. (2019, September 10). What's next for Registered Reports? Retrieved September 13, 2020, from https://www.nature.com/articles/d41586-019-02674-6

Chen, M. K., & Rohla, R. (2018). The effect of partisanship and political advertising on close family ties. *Science*, *360*(6392), 1020-1024.

Cohen, J. (1994). The earth is round (p<. 05). *American Psychologist*, *49*(12), 997.

Cooney, T. M., & Uhlenberg, P. (1992). Support from parents over the life course: The adult child's perspective. *Social Forces*, *71*(1), 63-84.

Cox, D. A. (2021, July 20). *The state of American friendship: Change, challenges, and loss*. The Survey Center on American Life. https://www.americansurveycenter.org/research/the-state-of-american-friendship-change-challenges-and-loss/

Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent–child similarity in intergroup attitudes. *Psychological Bulletin*, *139*(6), 1270.

Dickler, J. (2018, October 18). How parents pay for their kid’s college. Retrieved from https://www.cnbc.com/2018/10/17/how-parents-pay-for-their-kids-college.html

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology*, *1*(1), 263-292.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization?. *Public Opinion Quarterly*, *83*(1), 114-122.

Dunnett, C. W., & Gent, M. (1977). Significance testing to establish equivalence between treatments, with special reference to data in the form of 2 x 2 tables. *Biometrics*, 593-602.

Ellis, C., & Stimson, J. A. (2012). *Ideology in America*. Cambridge University Press.

Folley, A. (2020, June 25). Michigan candidate's daughter urges people not to vote for him in viral tweet. *The Hill*.

Fry, R., Passel, J. S., & Cohn, D. V. (2020). A majority of young adults in the US live with their parents for the first time since the Great Depression. Pew Research Center. https://www.pewresearch.org/fact-tank/2020/09/04/a-majority-of-young-adults-in-the-u-s-live-with-their-parents-for-the-first-time-since-the-great-depression/

Furman, W., & Buhrmester, D. (1985). Children's perceptions of the personal relationships in their social networks. *Developmental psychology*, *21*(6), 1016.

Goldenberg, A., Abruzzo, J., Willer, R., Halperin, E., & Gross, J. (2020, December 19). Homophily and Acrophily as Drivers of Political Segregation. https://doi.org/10.31219/osf.io/9xetr

Grant (2017, Aug. 11). Couples aren't the only ones overspending on weddings. *CNBC.* https://www.cnbc.com/2017/08/11/parents-often-overspend-on-wedding-expenses-for-adult-children.html#:~:text=Parents

Greene, S. (1999). Understanding party identification: A social identity approach. *Political Psychology*, *20*(2), 393–403.

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: ingroup favoritism enables discrimination. *American Psychologist*, *69*(7), 669.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology, 10*8, 553–561.

Greenwald, A. G., Gonzalez, R., Harris, R. J., & Guthrie, D. (1996). Effect sizes and p values: What should be reported and what should be replicated? *Psychophysiology*, *33*(2), 175–

183.

Guzzo, K. B., & Payne, K. K. (2018). Average age at first birth, 1970-2017. *Family Profiles, FP-18*, *25*.

Hoekstra, Rink, Morey, Richard, Rouder, Jeffrey N, & Wagenmakers, Eric-Jan. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, *21*(5), 1157-1164.

Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, *79*(1), 269-283.

Huddy, L., & Bankert, A. (2017). Political partisanship as a social identity. In *Oxford research encyclopedia of politics*.

Hyman, H. (1959). *Political socialization.* Free Press.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American journal of political science*, *59*(3), 690-707.

Iyengar, S., Konitzer, T., & Tedin, K. (2018). The home as a political fortress: Family agreement in an era of polarization. *The Journal of Politics*, *80*(4), 1326-1338.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129-146.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: a social identity perspective on polarization. *Public opinion quarterly*, *76*(3), 405-431.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336-352.

Jennings, M. K., & Niemi, R. G. (1968). The transmission of political values from parent to child. *The American Political Science Review, 62*(1), 169-184.

Jennings, M. K., Stoker, L., & Bowers, J. (2009). Politics across generations: Family transmission reexamined. *The Journal of Politics, 71*(3), 782-799.

Kalmoe, N. P., & Mason, L. (2019, January). Lethal mass partisanship: Prevalence, correlates, and electoral contingencies. In *National Capital Area Political Science Association American Politics Meeting*.

Kalmoe, N. P., & Mason, L. (2022). *Radical American Partisanship: Mapping Violent Hostility, Its Causes, and the Consequences for Democracy*. University of Chicago Press.

Kaplan, E., Spenkuch, J., & Sullivan, R. (2020). *Partisan spatial sorting in the United States: A theoretical and empirical overview*. Working Paper.

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. Personality and Social Psychology Review, 2(3), 196–217.

King, K. M., Feil, M. C., Halvorsen, M. A., Seldin, K., Shoda, Y., Lee, C., Smith, G (in press). Negative urgency as a state-level process. Accepted as a Stage 1 Registered Report by *Journal of Personality*. 10.17605/OSF.IO/9DBEJ

King, M. T. (2011). A point of minimal important difference (MID): a critique of terminology and methods. *Expert Review of Pharmacoeconomics & OutcomesRresearch*, *11*(2), 171-184.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*(5), 746-759.

Kriegbaum, K., Villarreal, B., Wu, V. C., & Heckhausen, J. (2016). Parents still matter: Patterns of shared agency with parents predict college students' academic motivation and achievement. *Motivation Science*, *2*(2), 97.

Krugman, P. R. (1991). *Geography and Trade*, Cambridge: MIT Press.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259-269.

Laursen, B. (1993). The perceived impact of conflict on adolescent relationships. *Merrill-Palmer Quarterly (1982-)*, 535-550.

Lemyre, L., & Smith, P. M. (1985). Intergroup discrimination and self-esteem in the minimal group paradigm. *Journal of Personality and Social Psychology*, *49*(3), 660.

Levendusky, M. S., & Stecula, D. A. (2021). *We Need to Talk: How Cross-Party Dialogue Reduces Affective Polarization*. Cambridge University Press.

Lykken, D. T. (1968). Statistical significance in psychological research. Psychological Bulletin, 70(3p1), 151.

Martin, E. (2019, Mar. 11). Here's how many millennials got money from their parents to buy their homes. *CNBC.* https://www.cnbc.com/2019/03/11/how-many-millennials-got-money-to-buy-homes-from-their-parents.html

Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487.

Meehl, P. (1990). Why summaries of research on psychological theories are often

uninterpretable. *Psychological Reports*, *66*, 195-244.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of science*, *34*(2), 103-115.

Milem, J. F. (1998). Attitude change in college students: Examining the effect of college peer groups and faculty normative groups. *The Journal of Higher Education*, *69*(2), 117-140.

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, *244*(4907), 933-938.

MIT Election Data and Science Lab. (2018). County presidential election returns 2000-2016. Harvard Dataverse. Retrieved from https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ

Mosleh, M., Martel, C., Eckles, D., & Rand, D. G. (2021). Shared partisanship dramatically increases social tie formation in a Twitter field experiment. *Proceedings of the National Academy of Sciences*, *118*(7).

Murphy, K. R., & Myors, B. (1999). Testing the hypothesis that treatments have negligible effects: Minimum-effect tests in the general linear model. *Journal of Applied Psychology*, *84*(2), 234.

Nadav-Greenberg, Limor, & Joslyn, Susan L. (2009). Uncertainty Forecasts Improve Decision Making Among Nonexperts. *Journal of Cognitive Engineering and Decision Making*, *3*(3), 209-227.

Newcomb, T. M. (1943). *Personality and social change; attitude formation in a student community*. Dryden Press.

Niemi, R. G., Danforth Ross, R., & Alexander, J. (1978). The similarity of political values of parents and college-age youths. *Public Opinion Quarterly*, *42*(4), 503-520.

Noller, P., & Callan, V. (1991). *The adolescent in the family*. Routledge.

Orben, A., & Lakens, D. (2019, May 30). Crud (Re)defined. https://doi.org/10.1177/2515245920917961

Oxford University Press, publisher. (2000). Oxford English dictionary. Oxford, England: Oxford University Press.

Park, A., Smart, C., Taylor, R., & Watkins, M. (2021). An Extremely Detailed Map of the 2020 Election. *New York Times*. Retrieved from https://www.nytimes.com/interactive/2021/upshot/2020-election-map.html

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922-934.

Pettigrew, T. F., Tropp, L. R., Wagner, U., & Christ, O. (2011). Recent advances in intergroup contact theory. *International Journal of Intercultural Relations*, *35*(3), 271-280.

Pew Research Center. (n.d.). *Religious landscape study.* *https://www.pewresearch.org/religion/religious-landscape-study/party-affiliation/republican-lean-rep/*

Pintea, S. (2010). The relevance of results in clinical research: Statistical, practical, and clinical significance. *Journal of Evidence-Based Psychotherapies*, *10*(1), 101.

Pizzolato, J. E., & Hicklen, S. (2011). Parent involvement: Investigating the parent-child relationship in millennial college students. *Journal of College Student Development*, *52*(6), 671-686.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, *112*(1), 160–164.

Robinson, R. (1993). Cost-utility analysis. *Bmj*, *307*(6908), 859-862.

Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6*(2), 88-106.

Rossiter, E. (2020). *The consequences of interparty conversation on outparty affect and stereotypes*. Working Paper. http://erossiter.com/files/conversations.pdf.

Santoro, E., & Broockman, D. (2021, September 23). The short-term, circumscribed, and conditional effects of cross-partisan conversation. https://doi.org/10.31219/osf.io/km3q2

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, *64*(6), 609.

Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, *15*(6), 657-680.

Schulz, R., Eden, J., & National Academies of Sciences, Engineering, and Medicine. (2016). Older adults who need caregiving and the family caregivers who help them. In *Families Caring for an Aging America*. National Academies Press (US).

Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good-enough principle. *American Psychologist. 40*(1), 73-83.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology*, *26*(6), 978–986. https://doi.org/10.1037/0012-

1649.26.6.978.

Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*(4), 267-281.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895.

Smiley, A. H., & Kaiser, C. R. (in prep). Across the aisle and over the fence: Communities, cross-party contact, and affective polarization.

Smiley, A. H., Glazier, J. J, & Shoda, Y. (under review). Large enough to matter, too small to be of consequence, or insufficient data? Towards a more meaningful statistical inference. https://psyarxiv.com/465ha

Standing, L., Sproule, R., & Khouzam, N. (1991). Empirical statistics: IV. Illustrating Meehl's sixth law of soft psychology: everything correlates with everything. Psychological Reports, 69(1), 123-126.

Standing, L., Sproule, R., & Khouzam, N. (1991). Empirical statistics: IV. Illustrating Meehl's sixth law of soft psychology: everything correlates with everything. *Psychological Reports*, *69*(1), 123-126.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96-103.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In *The social psychology of intergroup relations* (pp. 7–66). Monterey: Brooks Cole.

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods*, *6*(4), 371.

Umberson, D. (1992). Relationships between adult children and their parents: Psychological consequences for both generations. *Journal of Marriage and the Family*, 664-674.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, *70*(2), 129-133.

Webster, J., & Starbuck, W. H. (1988). Theory building in industrial and organizational psychology. *International Review of Industrial and Organizational Psychology*, *3*(1), 93-138.

West, E. A., & Iyengar, S. (2020). Partisanship as a social identity: Implications for polarization. *Political Behavior*, 1-32.

Westlake, W. J. (1976). Symmetrical confidence intervals for bioequivalence trials. *Biometrics*, 741-744.

Wojcieszak, M., & Warner, B. R. (2020). Can interparty contact reduce affective polarization? A systematic test of different forms of intergroup contact. *Political Communication*, *37*(6), 789-811.

**Acknowledgements**

I would like to thank my dissertation committee – Cheryl Kaiser, Sapna Cheryan, Yuichi Shoda, and Chris Parker – for their willingness to serve on my committee. In addition to supporting me through this expedited dissertation process, they all played a role in my development as a researcher and a scholar. I am especially thankful to Cheryl for her encouragement and timely feedback on my many drafts as I rushed to complete my dissertation much sooner than I had originally planned.

Thank you to my co-authors on this research – Jessica Glazier, Cheryl Kaiser, and Yuichi Shoda – for making this research possible. Research is a team sport. Thank you also to Jichan Kim and John Williams for assisting with recruitment at Liberty and Biola Universities. And a massive thank you to all of my research assistants for their hard work running participants. Special shoutout to Gabby Vanderlaan and Sonny Lim for helping to train other RAs and for coordinating a tricky study.

I would also like to thank my undergraduate mentors from Biola University, John Williams and Peter Hill, who inspired me to attend graduate school and helped me acquire the skills to hit the ground running on day one. I'd also like to thank Matt Fisher, who I interned with at Yale University when I was an undergraduate and he was a graduate student. We've since published two papers together, and both have faculty jobs… not too shabby! Thank you also to Jonathan Mercer for serving on my Generals Exam committee.

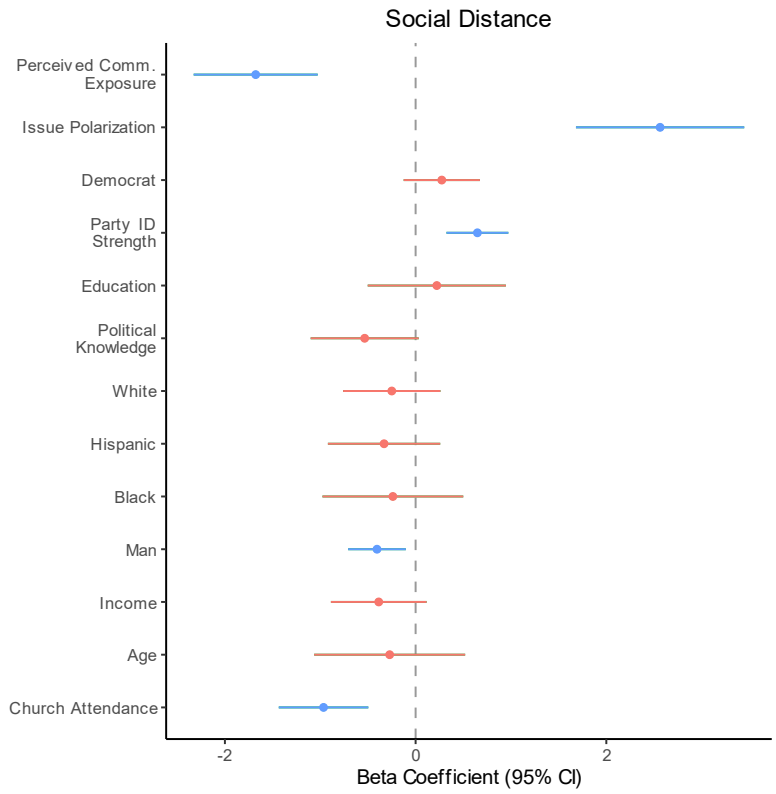
I'm thankful for the supportive environment from my classmates in the UW psychology department. Special thanks to my lab mates in the Social Identity Lab and the Shoda Lab. Thank you especially to Jon Gallegos, Jin Goh, Ishika Ray, and Justice Quame-Amaglo for your practical and moral support over the years.

And last-but-certainly-not-least, thank you to my family and friends. Thank you for teaching me, encouraging me, and supporting my education. To my friends from Biola who always called me an “artist” (because I was getting a BA in psychology while they got BS degrees in business)… you can address me as a philosopher now. Thank you especially to my parents who have continued to cheer me on (almost) as enthusiastically as they did at my soccer games. And to my brother, Kevin, who I forgot to acknowledge in my hooding ceremony speech… love you bro! Thank you also to my dogs, Annie and Miel. They have certainly made it harder to focus, slowed my productivity, and reduced the amount of sleep I get a night… But I couldn’t imagine going through a pandemic or a PhD program without them. And of course, thank you to my wife, Andrea, who has always believed in me. You give me strength, and I’m so excited for our next adventure – Nashville here we come!

# ARTICLE 1 APPENDICES

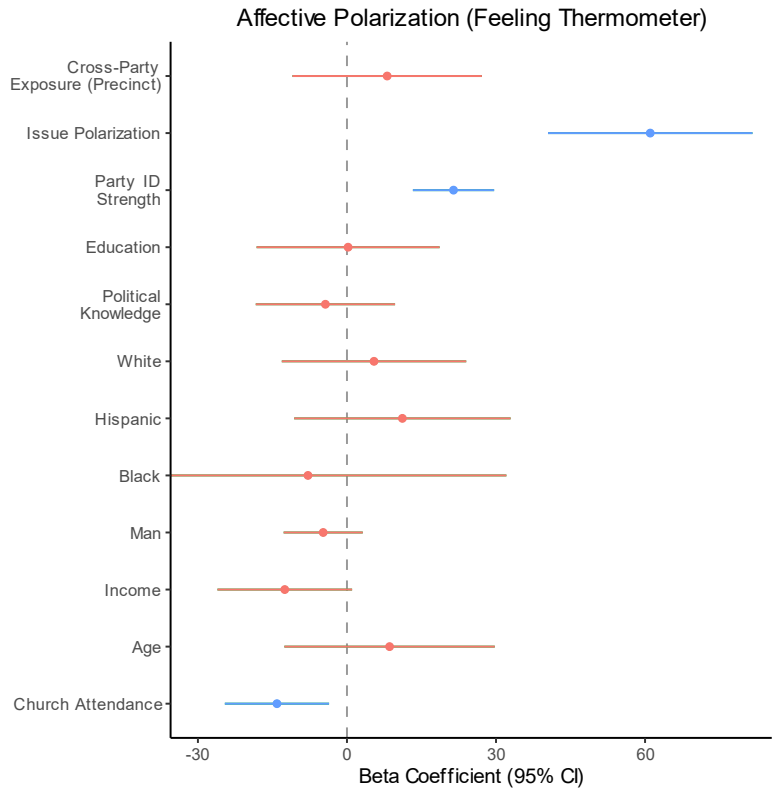
## Appendix S2

*Predicting Social Distance Using Perceived Cross-Party Exposure in Study 2.*



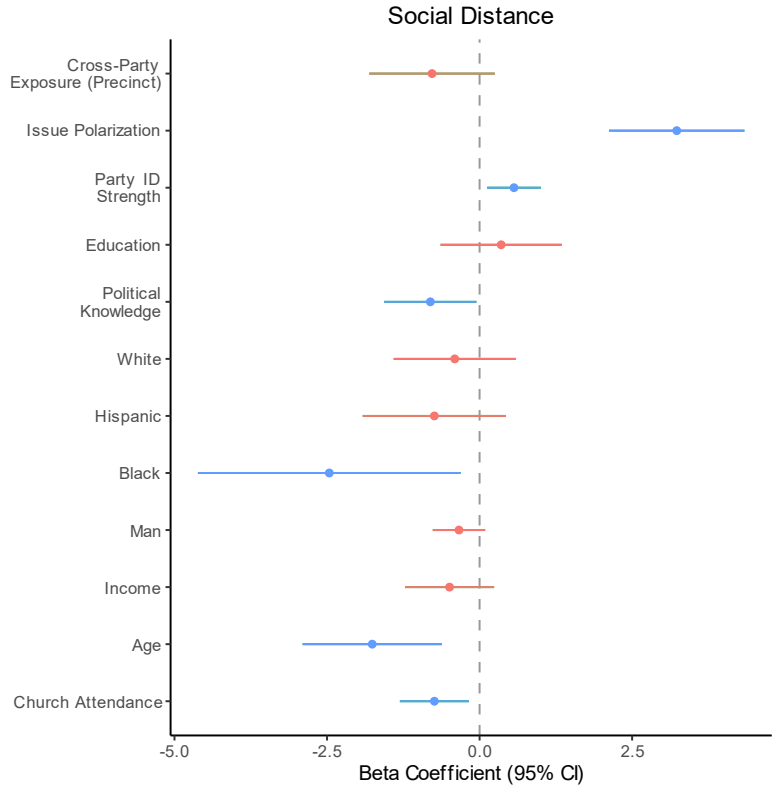

*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

*Predicting Affective Polarization (Towards Voters) for Republican Participants in Study 2*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

*Predicting Social Distance for Republican Participants in Study 2*




*Note.* Dots represent the point estimate for each beta coefficient, while whiskers (lines) represent the 95% CI. Items in blue are significant with $p < .05$, while items in orange are not statistically significant.

## Appendix S3

## UW Only Sample for Experiment

We recruited 23 Democrat-Republican dyads (46 participants) from UW, between April and November of 2020. One of these dyads was excluded for not following instructions (not having camera on during study), and two dyads were excluded because the participant who initially identified as a Republican identified as a Democrat when asked during the survey. One Republican participants' responses were excluded because they did not finish the survey (although they did complete the interaction tasks). In total, we included data from 20 Democrats ($M_{Age}$ = 18.60, SD = 1.31; 50% women, 50% men; 45% Asian/Asian-American, 35% White; 10% multiracial) and 19 Republicans ($M_{Age}$ = 18.67, SD = 1.37; 50% women, 50% men; 68% White; 32% Asian/Asian-American) from 20 dyads in the UW only version of this experiment. All UW participants were given course credit for participating in the study. We had pre-registered that we hoped to recruit 120 dyads from UW, but due to the incredibly small number of Republicans in our subject pool, we discontinued this version of the study to focus on recruiting participants from campuses with more Republican students.

*Pre-Intervention Means by Party in Study 3 (UW Only Sample)*

| Measure | UW Democrats | UW Republicans |
|---|---|---|
| Feeling Therm. AP (Voters) | 39.45 (25.52) | 16.63 (18.47) |
| Feeling Therm. AP (Party) | 32.95 (28.76) | 30.32 (22.01) |
| Social Distance (1-7) | 3.62 (1.76) | 2.42 (1.49) |

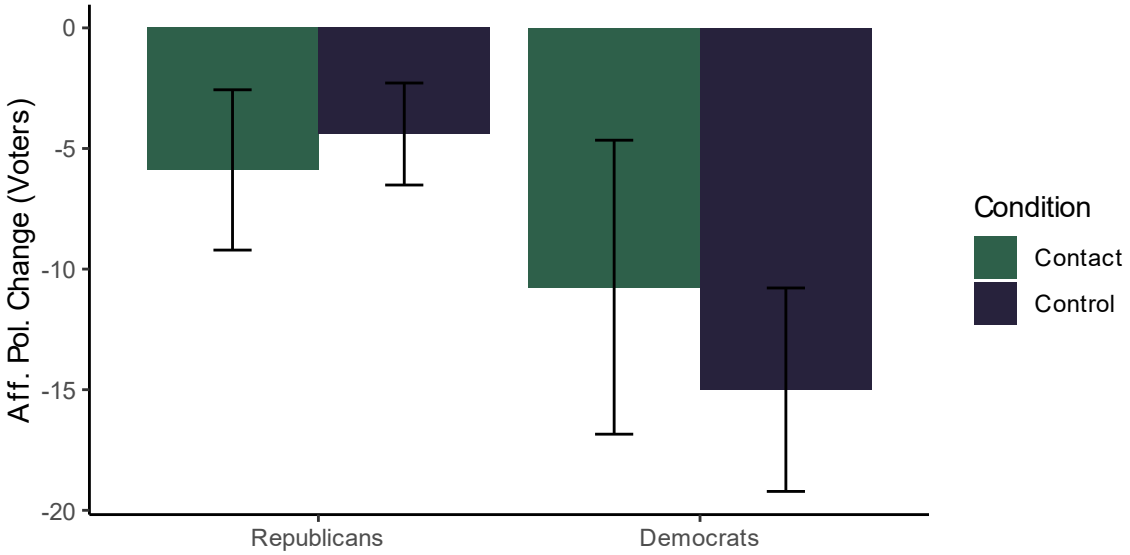
| | | |
|---|---|---|
| Partisan ID Index (0-1) | .55<br>(.17) | .50<br>(.16) |
| C-P Respect Important (1-7) | 6.70<br>(.47) | 6.47<br>(1.17) |
| C-P Friendships Positive (1-7) | 6.35<br>(.67) | 6.05<br>(1.03) |
| N | 20 | 19 |

In the UW only version of the study, there was no significant difference between the change in affective polarization towards party voters for participants in the contact condition ($M = -8.18$, $SD = 13.63$) compared with those in the control condition ($M = -10.18$, $SD = 12.65$), $t(33.0) = .47$, $p = .64$, 95% CI = [−6.67, 10.69] (see Fig. below). The difference between conditions in the amount of change in affective polarization regarding the parties as a whole (not voters) was approaching significance, with a marginally larger reduction in the contact condition ($M = -11.06$, $SD = 17.09$) compared to the control condition ($M = -1.59$, $SD = 12.09$), $t(27.6) = -1.94$, $p = .06$, 95% CI = [−19.47, .53] (see Fig. below). Additionally, there was no significant difference in pre-post intervention differences for social distance between the contact condition ($M = -.39$, $SD = .75$) than for participants in the control condition ($M = -.37$, $SD = 1.13$), $t(34.8) = -0.09$, $p = .93$, 95% CI = [−.65, .59] (see Fig. below). In the UW only version, 66% of the (nine) dyads in the contact condition were successful and in the control version 92% of the (12) dyads were successful.

Although in our cross-campus version of the study we observed meaningful reductions of affective polarization and social distance for those in the contact condition, we did not observe the same pattern in the UW only version of the study. We found no significant effect of contact on affective polarization with regards to voters, or social distance, while contact appeared to

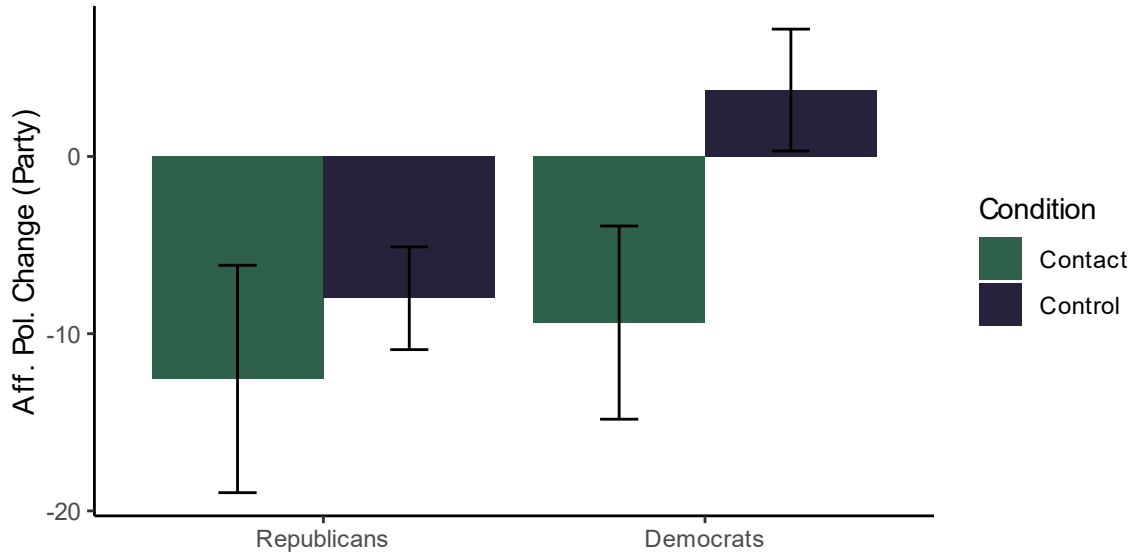
marginally reduce affective polarization towards parties as a whole. However, the sample size for this version of the study was quite small (20 dyads). Additionally, Republicans at the University of Washington have a large degree of cross-party exposure to begin with, which may make additional cross-party contact less meaningful. Still, the mixed findings from this study suggest that future research is needed to examine the causal relationship between cross-party contact and affective polarization.

*Voter-Directed Affective Polarization Change by Party and Condition in UW-Only Study*



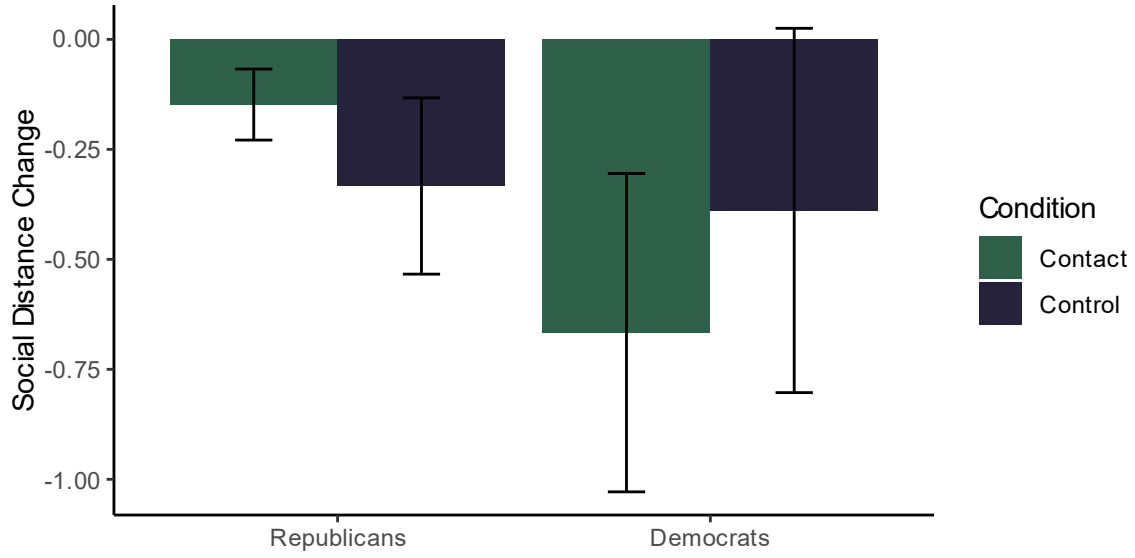

*Note.* Error bars denote SE.

*Party-Directed Affective Polarization Change by Party and Condition in UW-Only Study*




*Note.* Error bars denote SE.

*Social Distance Change by Party and Condition in UW-Only Study*



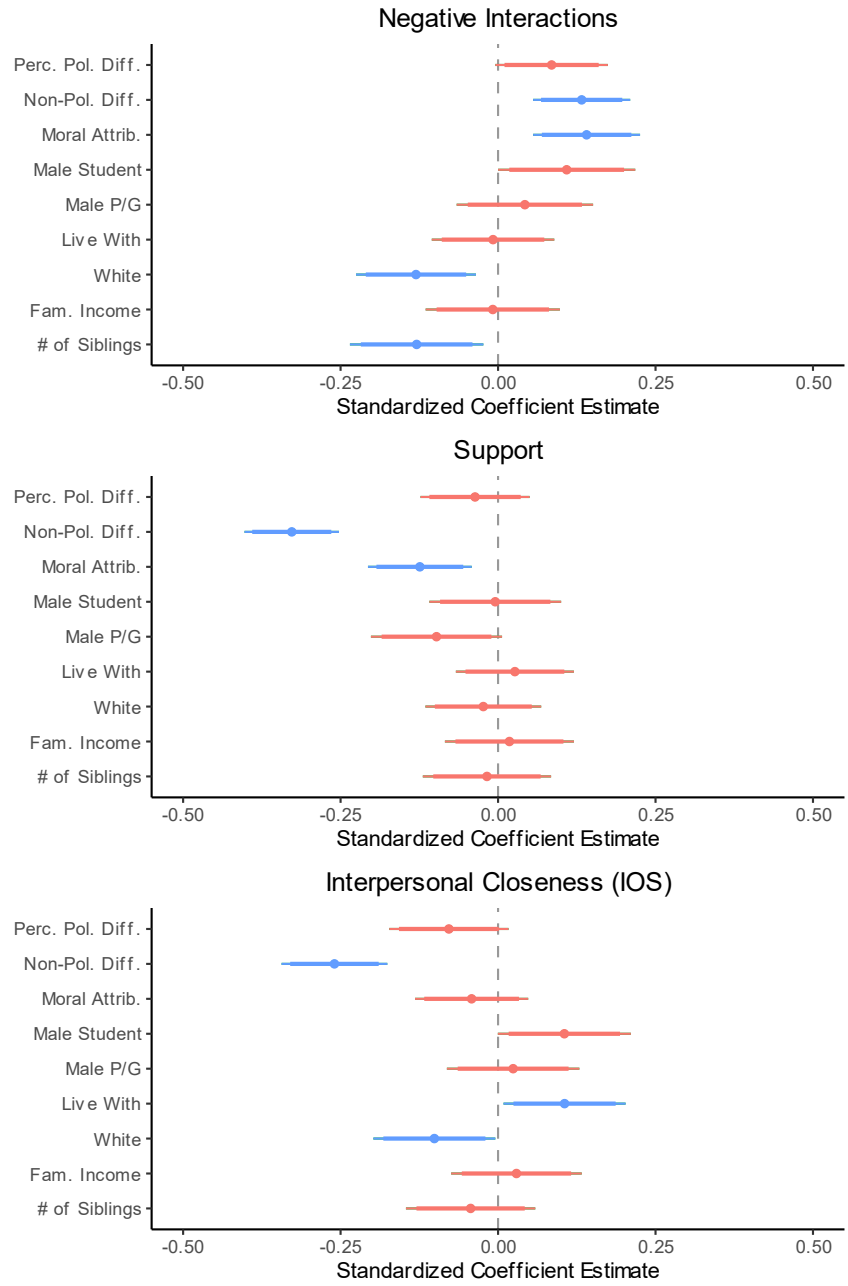

*Note.* Error bars denote SE.

# ARTICLE 2 APPENDICES

## Appendix S2

## Study 2 Pre-Registered Models (No Deviations from Analysis Plan)

*Fixed-Effect Predictors of Student-Parent Relationship Quality in Study 2*




*Note*. Standardized β coefficients of fixed effects for three levels of relationship quality. Models had random intercepts for each student-parent dyad. Responses are from both parents and

students. Points = estimates. Thick whiskers = 90% confidence intervals. Thin whiskers = 95% confidence intervals. Estimates significantly different than 0 with $\alpha = .05$ are shown in blue, while non-significant effects are shown in orange. "P/G" refers to parent or guardian.

## Exploratory Analyses from Study 2

Responses to questions about shame and pride showed clear floor and ceiling effects. On a 1 to 7 scale with high scores representing more shame in their child, the mean score from parents was 1.32 (*SD* = .94), while student perceptions of how ashamed their parents are of them were just slightly higher (*M* = 1.87, *SD* = 1.16). On average, parents indicated that they were very proud of their children (1-7 scale, *M* =6.61, *SD* = .88) while students perceived their parents were proud of them as well (*M* = 6.19, *SD* = 1.32). For parents, there is no significant zero-order correlations between perceived political difference and being proud of their child ($r = -.16$, 95% CI = [−.19, .05], $p < .23$) while the relationship between perceived political difference and shame was approaching significance ($r = .12$, 95% CI = [−.003, .24], $p = .06$). For students, when they perceived their political views were more different, they perceived significantly less pride from their parents ($r = -.15$, 95% CI = [−.26, −.03], $p = .01$) and significantly more shame ($r = .17$, 95% CI = [.06, .29], $p = .004$).

We examined open-ended responses about whether students and parents avoid particular political discussions. Overall, most respondents reported that they did not intentionally avoid talking about political issues. Yet, there were many students and parents who did avoid particular topics or discussing politics altogether. Many simply wanted to avoid arguments, including one mother who said that she avoids particular topics because she "want[s] us to live peacefully with each other." Some respondents wanted to protect their family members feelings (e.g., "Yes because she always ends up angry and often in tears") while others did so in order to prevent

their family member from being upset with them (e.g., “I am sometimes afraid that my child will think differently of me because of my political views.”). For others, the overlap between specific political topics and their own social identities made them particularly avoidant of discussing those issues. One student wrote, “I avoid discussing LGBTQ rights because I am a closeted queer-identifying person and I don't want to potentially out myself or have to face the fact that my parents are not supportive of the LGBTQ community.”

When asked if they end political conversations when they get too emotional, most respondents said that they either did not have emotional conversations about politics or that they do not end conversations when emotions ran high. Some were even intentional about allowing emotional conversations, including one mother who wrote, “No, I feel it is important to have a free-flowing conversation with one another and not walk away in a fight.” However, many indicated that they practice conflict aversion and end political conversations if they get too intense. Some expressed that politics were not worth harming their relationship (e.g., “Politics are not worth getting in arguments with your child about.”). One student responded that she had ended conversations often, writing “My mother will easily switch to shutting me down without much thought and these conversations are usually ones that I'm really passionate about and that really affect my life, so I will shut down and quit wanting to talk to her about it once she gets angry and pushy with me.” One respondent even responded that often his “wife will need to intervene and tell us to cool off.” Does the practice of conflict avoidance or aversion help to protect the relationship? Or does it lead to greater distance between students and their parents? In Study 3 we build upon these open-ended responses by quantifying the degree of avoidance and aversion in order to explore how these factors relate to relationship quality.

Participants were also asked open ended questions about how politics affect their

relationships, if their family member has any political views that cause them to think less of them, and if their family member holds any political views that are in opposition to any of their social identities. Consistent with our quantitative findings, most respondents suggested that political differences did not harm their relationships. This was especially so for parents – only a single parent indicated that one of their child's political views caused them to think less of their child. The vast majority of students responded similarly. However, one clear pattern emerged: many LGBTQ+ students reported that their parents' views on issues of gender and sexuality were harmful to their relationship and to their own mental health. One student wrote, "My mother is a firm believer that there are only two genders (male and female) and that marriage is only valid between a man and a woman. As someone currently exploring my gender identity and quite sure of my attraction in the same-sex, these views do make me feel hurt, angry, and unsafe." Several other LGBTQ+ students reported a fear of coming out to their parents because of their political views (e.g., "Although she doesn't know that [I am gay], it does make me feel worried about the future and whether I'll ever be able to tell her without feeling ashamed."). Although generally we find little support for political differences predicting worse relationship quality in general, these findings suggest that political disagreements do have the potential to meaningfully harm relationships, particularly when political views intersect with important social identities like sexual orientation and gender identity.

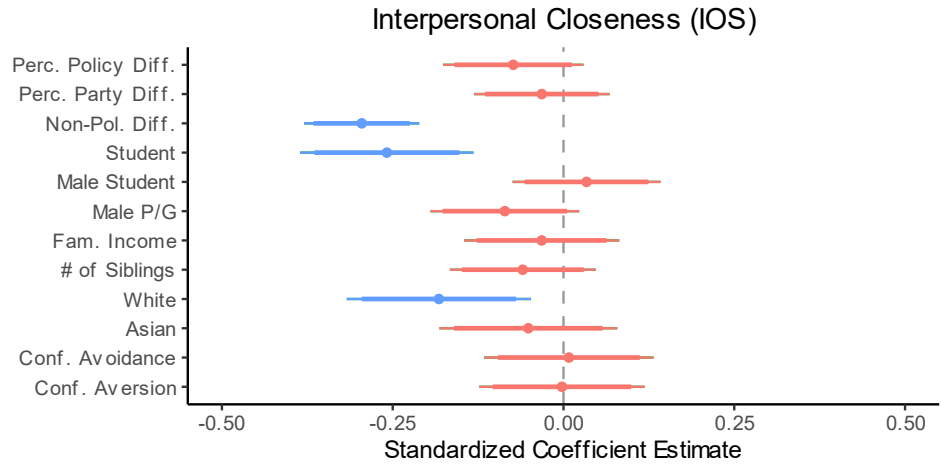
An extensive body of research suggests that parental support is important for the mental health of LGBTQ+ adolescents, with rejection being associated with particularly negative mental health outcomes (e.g., Mills-Koonce et al., 2018; Ryan et al., 2009, 2010). Open ended responses suggest that while political views may be one way in which parents signal their support for LGBTQ+ individuals, which can have immense consequences for their children. Still, more

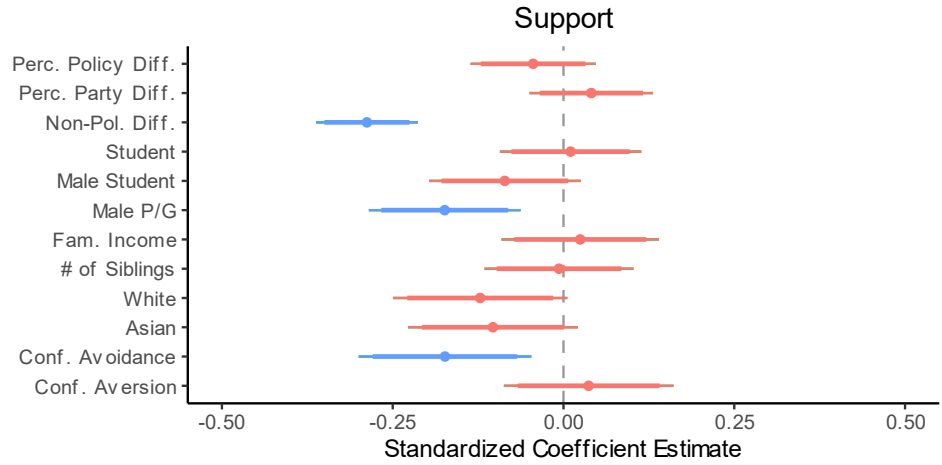
research is needed to examine how parents' policy views and political partisanship can signal attitudes towards sexual minorities.

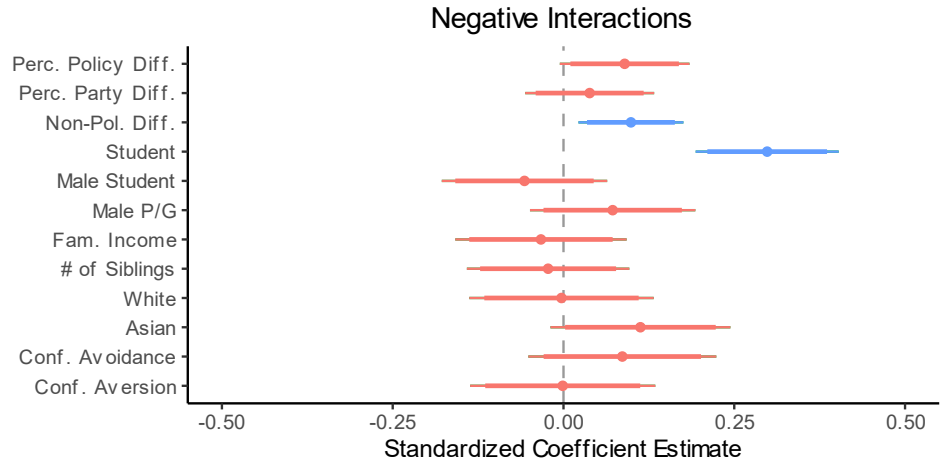
## Appendix S3

## Study 3 Pre-Registered Models (No Deviations from Analysis Plan)

*Fixed-Effect Predictors of Student-Parent Relationship Quality in Study 3*



*Note*. Standardized β coefficients of fixed effects for three levels of relationship quality. Models had random intercepts for each student-parent dyad. Responses are from both parents and students. Points = estimates. Thick whiskers = 90% confidence intervals. Thin whiskers = 95% confidence intervals. Estimates significantly different than 0 with α = .05 are shown in blue, while non-significant effects are shown in orange. "P/G" refers to parent or guardian.

## Exploratory Analyses from Study 3

The majority (65.4%) of students and parents identified with the same political party (see Table X). We also scaled partisan identity to a 0-1 scale (0 = Republican, 0.5 = independent, 1 = Democrat) and ran a correlation, which revealed a Pearson's correlation of $r$ = .45. These levels of parent-child correspondence are similar to those observed by previous studies of late adolescents and their parents (Jennings & Niemi, 1968; Jennings et al., 2009).

**Table X**

*Partisanship of Parents and their Children (College Students) in Study 3*

| Parent Party | Student Party | | | | |
|---|---|---|---|---|---|
| | Republican | Independent | Democrat | *Total* | Same Party |
| Republican | **12** | 16 | 15 | 43 | 27.9% |
| Independent | 4 | **28** | 20 | 52 | 53.8% |
| Democrat | 4 | 22 | **113** | 139 | 81.2% |
| *Total* | 20 | 66 | 159 | | |

*Note.* Cells where students and parents identify with the same party are bolded.

We explored how accurate students and parents were in their perceptions of their family members partisan identification. To do this, we compared the party that each respondent identified with to the party their family member thought they would identify as. Parents correctly perceived their child's party identification in a majority (76.3%) of cases. Students' correctly

perceived their parents partisanship in 83.7% of cases, which is marginally more accurate than parent perceptions, $\chi^2 = 3.66$, $p = .06$.

Parents' accuracy in perceiving their child's partisanship varied based on the student's actual partisanship, where students who identified as Democrats were correctly perceived 87.8% of the time but independents (54.4%) and Republicans (65.0%) were correctly identified much less frequently (see Table X). Likewise, student perceptions of their parent's partisanship were highly accurate when their parent identified as a Democrat (92.1%) or Republican (86.0%) but relatively lower when they identified as an independent (59.6%; see Table X).

**Table X**

*Student Partisan Identity and Parents' Perceptions in Study 3*

| Student Party | Student Perceived Party (by Parents) | | | | |
|---|---|---|---|---|---|
| | Republican | Independent | Democrat | *Total* | Correctly Perceived |
| Republican | **13** | 6 | 1 | 20 | 65.0% |
| Independent | 6 | **37** | 25 | 68 | 54.4% |
| Democrat | 4 | 14 | **128** | 146 | 87.8% |

*Note.* Correct perceptions are bolded.

**Table X**

*Parent Partisan Identity and Students' Perceptions in Study 3*

| Parent Party | Parent Perceived Party (by Student) | | | | |
|---|---|---|---|---|---|
| | Republican | Independent | Democrat | *Total* | Correctly Perceived |
| Republican | **37** | 2 | 4 | 43 | 86.0% |
| Independent | 8 | **31** | 13 | 52 | 59.6% |
| Democrat | 2 | 9 | **128** | 139 | 92.1% |

*Note.* Correct perceptions are bolded.

# ARTICLE 3 APPENDICES

## Appendix S1

### The Rationale for using Confidence Intervals to conduct MET and EqT

MET and EqT require estimating the probability of obtaining the observed value (or more extreme (MET) or less extreme (EqT)) of a statistic of interest (e.g., differences in means, correlations, proportions), randomly drawn from a population in which the null hypothesis $H'_0$ holds. Unlike the null hypothesis ($H_0$) of a traditional NHST, MET and EqT use the TOST (two one-sided significance testing), which typically assumes a certain, pre-specified, non-zero value in the population as null hypotheses ($H'_0$s).

For example, one-tailed METs with the alpha level of X% require determining the 100-X percentile (e.g., $95^{th}$ percentile) of the sampling distribution of the statistic under the null hypothesis $H'_0$ that the population value is a certain hypothesized non-zero value. Murphy and Myors (1999) derived such sampling distributions for the F statistic. In order to make MET applicable to a wide variety of statistics, however, the present approach makes use of the fact that the sampling distribution under $H'_0$ is already taken into account when determining confidence intervals. Thus, as described below, we propose that whenever confidence intervals can be determined, it is possible to "reverse engineer" the critical values for use in MET.

We start by first reviewing the logic for estimating the confidence interval of a quantitative variable:

(1) Suppose we draw many samples of the same size from a population for which the statistic of interest is $H'_0$. (Note we are using $H'_0$ to refer also to the population value of the statistic under the null hypothesis $H'_0$). Let $\delta_{95}$ be the distance between $H'_0$ and

the 95th percentile of the sampling distribution from this population, and $\delta_5$ be the distance between $H'_0$ and the 5th percentile of the sampling distribution. Thus, the 95$^{th}$ percentile of this distribution is $H'_0 + \delta_{95}$, and the 5$^{th}$ percentile is $H'_0 - \delta_5$.

(2) Let $\bar{x}_i$ be the observed statistic of interest in sample *i* drawn from this population (e.g., the sample mean). Then by definition, Prob ($\bar{x}_i < H'_0 - \delta_5$) = 5%. It also follows that Prob ($\bar{x}_i + \delta_5 < H'_0$) = 5%.

(3) Let $\delta_i{}^*_5$ be an unbiased estimate of $\delta_5$ based on the data in sample *i*. If we can assume, at least approximately, that Prob ($\delta_5 < \delta_i{}^*_5$) = Prob ($\delta_5 > \delta_i{}^*_5$) = 0.5, then because Prob ($\bar{x}_i + \delta_5 < H'_0$) = 5%, Prob ($\bar{x}_i + \delta_i{}^*_5 < H'_0$) = 5% in the long run. If we define $\bar{x}_i + \delta_i{}^*_5$ as the upper bound of the confidence interval based on sample *i*, $CI_{upper\_i}$, then Prob ($CI_{upper\_i} < H'_0$) = 5% in the long run.

That is, the probability that the upper bound of the CI is below $H'_0$ is 5%. In other words, the probability that the entire CI is below $H'_0$ is 5%.

(4) Similarly, let $\delta_i{}^*_{95}$ be an unbiased estimate of $\delta_{95}$ based on the data in sample *i*. Then because from the definition of $\delta_{95}$, Prob ($H'_0 + \delta_{95} < \bar{x}_i$ ) = 5%. Thus, in the long run, Prob ($H'_0 + \delta_i{}^*_{95} < \bar{x}_i$ ) = Prob ($H'_0 < \bar{x}_I - \delta_i{}^*_{95}$ ) = 5%. If we define $\bar{x}_i - \delta_i{}^*_{95}$ as the lower bound of the confidence interval based on sample *i*, $CI_{lower\_i}$ ,then Prob ($H'_0 < CI_{lower\_i}$) = 5% in the long run. That is, the probability that the entire CI is above $H'_0$

is 5%.

(5) Thus, it makes sense to define the CI as $[\bar{x}_i - \delta_i{}^*_{95}, \bar{x}_i + \delta_i{}^*_5]$. With this definition of the CI, the CI computed in each sample will be entirely above $H'_0$ in 5% of the samples randomly drawn from the population where $H'_0$ is true. Similarly, the CI will be entirely below $H'_0$ in 5% of the samples.

As a result, the probability that the CI will contain the population value is 90%.

Now let us reverse engineer $\delta_i{}^*_5$ and $\delta_i{}^*_{95}$. Suppose for a statistic of interest, there is a well-accepted formula for determining the CI based on a given sample. Then, if we define, as described above, $CI_{upper\_i} = \bar{x}_i + \delta_i{}^*_5$ and $CI_{lower\_i} = \bar{x}_i - \delta_i{}^*_{95}$, where $CI_{upper\_i}$ and $CI_{lower\_i}$ refer to the upper and lower bounds of the CI by applying the well-accepted formula to the data in sample *i*, we can "reverse engineer" $\delta_i{}^*_5$ and $\delta_i{}^*_{95}$ as follows:

$$\delta_i{}^*_5 = CI_{upper\_i} - \bar{x}_i$$

and

$$\delta_i{}^*_{95} = \bar{x}_i - CI_{lower\_i}$$

By the definition of MET as described in the main text, a one-tailed MET is considered significant when $\bar{x}_i > NR_U + \delta_{95}$. But of course, $\delta_{95}$ is unknowable empirically. However, as shown above, in the long run, $\bar{x}_i > NR_U + \delta_{95}$ when $\bar{x}_i > NR_U + \delta_i{}^*_{95}$, or when $\bar{x}_i - \delta_i{}^*_{95} > NR_U$. That is, a one-tailed MET is significant when $CI_{lower\_i} > NR_U$ (i.e., the entire CI is above $NR_U$).

Now let's us consider a two-tailed MET. A two tailed test is significant when $\bar{x}_i > NR_U + \delta_{95}$ or $\bar{x}_i < NR_L - \delta_5$. As shown above, the first condition is met when $CI_{lower\_i} > NR_U$ (i.e., the entire CI is above $NR_U$). It can be shown similarly that the latter condition is met when $CI_{upper\_i} < NR_L$ (i.e., the entire CI is below $NR_L$). In other words, a two-tailed MET is significant when the entire CI is either above $NR_U$ or when the entire CI is below $NR_L$.

Note that when the sampling distribution is symmetrical, the above becomes simplified because $\delta_5 = \delta_{95}$. But we did not make that assumption so that our reasoning applies even to statistics for which the sampling distribution is not symmetric.

## Appendix S2

## Peer-Review

An important goal of MET is to encourage a discourse on the magnitude of effects. We believe that a peer-review process will help achieve this goal, while also promoting greater transparency of how authors arrive at the null region for use in MET.

To achieve these goals, it is important to consider the incentives that may influence researchers while setting $H'_0$. For example, researchers may be motivated to set a "low bar" with null regions very close to zero out of fear that setting the null region to be too large would make it difficult to reject the MET null hypothesis.

There are multiple points at which $NR_L$ and $NR_U$ values can be reviewed and refined in the research/publication process. First, a researcher preregistering their studies and/or analyses can request peer reviews of the proposed $H'_0$ from colleagues knowledgeable about the domain of research. The names of those who review the $H'_0$, along with their reviews and the $H'_0$ they suggested, should be included as part of a pre-registration and in the eventual publication (with their consent). Both the authors and the reviewers should be mindful that an unreasonably small or large $H'_0$ may call their judgment into question.

Second, journal editors could request the reviewers of a manuscript reporting MET to specifically comment on the appropriateness of the $H'_0$ chosen by the authors. If the editor and/or reviewers determines that the $H'_0$ is inappropriate, then the editor can ask the authors to repeat

their MET with a more appropriate value for $H'_0$. In this context, manuscript reviewers who suggest an unreasonably large $H'_0$ need to consider that their own future (or past) work in the same domain would be held to the same standards.

**Sample Review Request**

To help facilitate the process of requesting an informal peer-review, below is an example of a request a researcher may send to another researcher, requesting peer-review of their $H'_0$ values prior to data analysis. A request by editors to reviewers of a submitted manuscript may be generated by editing this sample request.

Hello Dr. Garcia,

We are writing you in order to request your peer-approval of the null hypothesis we have chosen for our study. Because of your expertise in technology and learning in public middle schools in large metropolitan areas, we believe that you would be able to effectively evaluate whether our null hypothesis is reasonable.

Our study examines the effect of giving all middle school students their own tablet. We plan to evaluate the utility of doing so by comparing schools that have recently implemented a program that gives students tablets to schools that have not, and assessing students on a standardized math exam at the end of the year.

Because of the great expense of implementing a tablet program, we used utility analysis to determine the size of the effect below which it would be not worthwhile to implement the proposed program. Specifically, we have decided that a 5% increase in the

proportion of students who meet the grade level standard for math is appropriate as the smallest effect of consequence. This will be used as a basis for statistical tests (please see the attached article describing them) determining if the true comparative effectiveness of the free tablet program is large enough to warrant further consideration, because other interventions such as extra tutoring have been shown to result in at least a 5% increase for the same cost.

If you respond to this message with feedback, your name and a summary of your response will be included in a supplement of our paper. Please let us know if you agree that this is a reasonable basis for specifying the smallest effect of consequence. If you do not agree, we would appreciate any feedback you have on how to make this value more appropriate and/or a suggestion of a better value for consideration as the smallest effect of consequence. We value your expertise and believe that your feedback will improve the quality of our statistical inference.

Sincerely,

Adam Smiley, Jessica Glazier, and Yuichi Shoda

**Searchable H'$_0$ Database**

Initially practices such as those above will rely on the scientific judgment of authors, reviewers and editors. However, we believe it should be possible for the scientific community to start a searchable database of H'$_0$s and accumulate H'$_0$s that have been judged reasonable (along with the names of those who provided the judgment) for a particular research question, research

goal, and participant populations. The database will also serve as a repository of H'$_0$ currently considered the minimum for a meaningful result for each relevant context (e.g., research question and goals, participant populations, etc.), along with a record of how often empirical research has rejected them.

As each scientific discipline strives toward establishing a useful database of H'$_{0s,}$ we believe that the public discourse about the H'$_0$s itself will encourage transparency and communicate the meaning of the observed effects, beyond the question of whether or not it can arise if there was absolutely no true effect. Though the details for the creation of such a database is beyond the scope of this paper, we believe that the scientific community can create and maintain an effective database (or multiple databases).

**Avoiding Conflict of Interest**

As with peer-review of journal articles, those chosen to review the H'$_0$ should avoid conflicts of interest. Differences between superiors and subordinates in the power they hold over the other should be considered when determining who should be asked to review the $NR_L$ and $NR_U$ values, as a subordinate might feel pressure accept the value. For example, it would be a conflict of interest if a graduate student were asked to approve of a H'$_0$ value proposed by a faculty member in their own graduate program, or if an applicant on the academic job market were asked to review a H'$_0$ value for a faculty member from a university they are seeking to gain employment from. Avoiding conflicts of interest is important for the quality of research being conducted, as researchers should be able to review proposed H'$_0$s without experiencing undue pressure. To this end, researchers should use their best judgment when considering potential reviewers in order to avoid such conflicts of interest, and practice transparency and

accountability, for example by including a statement declaring no conflict of interest, and disclosing all potential sources of conflict.