Supplementary Information: Cross-Partisan Conversation Reduced Affective Polarization for Republicans and Democrats Even After the Contentious 2020 Election

December 21, 2022

Contents

Appendix A   Ethical Considerations

The human subjects research conducted for our study adheres to APSA's Principles and Guidance for Human Subjects Research. The study was approved prior to data collection by the authors' institutional review board under expedited review. The IRB deemed: "the criteria for approval are met per 45 CFR 46.111 and/or 21 CFR 56.111 as applicable. Project determined to be minimal risk per 45 CFR 46.102(i) and/or 21 CFR 56.102(i) as applicable." A modification to the study protocol was also approved in which the only change was to increase the number of participants enrolled in the study.

All participants who completed the pre-treatment survey were paid $0.75. All participants who were invited to return to the conversation/short essay portion of the survey were paid $2.00 if they returned, completed their task, and completed the post-treatment survey that immediately followed the task. Participants who returned but could not complete the task because their conversation partner did not return were still paid the full $2.00 compensation. Additionally, participants were incentivized to have "thoughtful, thorough" participation with a bonus of $1.00 for doing so. Without the bonus, participants earned $2.75 for approximately 12 minutes of work, resulting in an hourly wage of $13.75. Finally, participants who completed the study were invited back to a follow-up survey asking one question that took only seconds to answer. All participants who completed this question were paid $0.50. All participants self-reported living in the United States and were paid above federal minimum wage.

Prior to beginning the pre-treatment survey, all participants read an information sheet to obtain their informed and voluntary consent. Participants had to check a box indicating "I consent" to proceed, which is how we documented consent in an online environment. Participants also had the option to select "I do not consent to participate," which would allow them to leave the survey immediately.

We will share the full consent information sheet upon request. Importantly, among other things, our information sheet shared how long we anticipated the tasks would take, the payment for completing the tasks, and how they would be paid. We also told participants that they could choose to stop participating at any time. We also told participants that we would keep the information they provided confidential, and that we would not be collecting any personally identifying information.

Finally, we told participants that anonymized transcripts of their conversation or short essay would be made available for research purposes as described, involving removing any people's names, places, religious or cultural backgrounds, occupations, family relationships, and any other potentially identifying information that they may have disclosed in their conversations (even though they were explicitly not prompted to do so).

Our participants came from Amazon's Mechanical Turk, which has a subject pool broadly demographically diverse within the U.S. population and not comprised mainly of members of groups we should consider vulnerable or marginalized. This research did not differentially harm particular demographic groups.

Appendix B   Sample Sizes and Attrition

3,483 participants completed an initial pre-treatment survey, and of them, 1,801 agreed to return to a follow up task, described as having a conversation with another MTurk Worker or writing a short essay. From those who agreed to return, only 1,032 were randomly assigned to treatment and invited back. Because our design required balance on partisanship, but MTurk has more Democrats than Republicans, the number of participants included in the design was smaller than the number that agreed to participate.

Of the 1,032 participants included in the design, 698 (67%) completed the conversation or short essay and the post-treatment survey. However, because treatment is assigned at the *partnership* level, we pre-registered dropping all cases where the full partnership does not complete the task, resulting in our main analyses having a sample size of 578 participants with an equal number of Democrats (289) and Republicans (289) since each partnership has one Democrat and one Republican by design.

410 of the 578 participants (70.9%) included in the main analyses returned three days later to complete the follow-up to assess durability of treatment effects. We preregistered analyzing all participants that complete this survey item, regardless of whether their partner also completed the item.

We next assess whether rates of attrition differ across treatment and control groups. The first model in Table A.1 shows that, of all partnerships *in the design* (N=516 partnerships), a partnership was no more or less likely to complete the task depending on if they were assigned to treatment or control. This means we didn't have differential attrition for those assigned to have a conversation, a task requiring more coordination than a short essay.

Furthermore, the second model in Table A.1 shows that an individual is just as likely to return to take the follow-up survey whether they were a part of a partnership that was assigned to write short essays or have a conversation. Of the 578 individuals invited back to complete the follow up, 410 did, and it was no more or less likely for an individual to complete this task if they completed a conversation, for example.

Finally, the third model in Table A.1 shows that change in outparty affect, our main outcome of interest, does not explain who returns to take the follow-up survey. Therefore, it is unlikely that the follow-up survey consists of only respondents who had strong treatment effects.

Table A.1: Assessing Differential Rates of Attrition

	Partnership Completed Post-Treatment Survey	*Individual* Completed Follow-up Survey	
(Intercept)	0.55*	0.71*	0.71*
	(0.03)	(0.03)	(0.03)
Conversation	0.02	−0.01	−0.02
	(0.04)	(0.04)	(0.04)
Change in Outparty Animosity			0.00
			(0.00)
N	516	578	578

Note: * p < 0.05. First model tests whether treatment assignment explains whether the partnership finishes the task. The second and third models test whether treatment assignment and effect of treatment explains whether participant returns for follow-up survey three days later.

Furthermore, in Appendix C, we show that we have balance on all pre-treatment observables except one in the sample used for analyses after attrition. We fail to find any strong determinants of attrition from these balance tests.

Appendix C Blocking Strategy

For our experimental design, we needed to construct partnerships from participants in our sample and then randomly assign each partnership to either the control condition of writing individual short essays (and thus not interacting with their assigned partner) or the treatment condition of conversation with their assigned partner. Because we were randomly assigning treatment at the partnership level, this is usually called "cluster" randomization, whereby clusters (here, partnerships) are the unit of randomization and all units within the cluster are assigned to the same condition.

To improve balance across conditions, and thus improve precision of our estimates and power of our tests, we also took the opportunity to implement a block randomized design. A block randomized design finds units of randomization (here, partnerships) that will have similar potential outcomes, and randomly assigns treatment within each block. In our experiment each block contained two partnerships, where one partnership was randomly assigned to have a conversation (treatment) and the other was assigned to write individual short essays (control).

In addition to randomly assigning treatment within each block, we also needed to randomly assign the partnerships themselves. We constructed blocked partnerships that would maximize partnership-level similarity. In other words, we created two partnerships that would be good counterfactuals for each other, placed them into a block, and then randomly assigned one of those partnerships to treatment and one to control via block randomization.

For example, consider a block to have partnerships A and B. If partnership A in a block consisted of a younger Republican and an older Democrat, partnership B ought to have a younger Republican and an older Democrat as well to create a good counterfactual. However, the people within a partnership do not have to be different according to this blocking strategy. If partnership A in a block has two young women (one Republican and one Democrat by design), then partnership B ought to have two young women as well (again, one Republican and one Democrat by design).

Across blocks, how different the people within a partnership are varies. As in the example above, some people will be very similar to their partner and some will be very different. This is because we randomly assign partnerships. However, *within* a block, the ways in which the people in a partnership are different will look similar in the two partnerships.

This design allows us to be confident we've created good counterfactuals for conversation, helps

us achieve balance on important covariates at the partnership level across experimental conditions, and thus improves precision of our estimates and power of our tests.

More specifically, we create blocked partnerships using the following covariates from our pre-treatment survey: age, education, non-white, non-male, partisan identity strength, moralization of partisanship, political interest, vote choice, preference for outparty conversation, feelings about the 2020 election (fear, hope, anger, pride), perceptions of Trump and Biden's conduct since the election, attention paid to the 2020 election, political status of parties, feeling like a "winner" after the election, social polarization, perceptions of election integrity, and feeling thermometers toward Trump, Biden, Republicans, and Democrats.

Figure A.1 shows balance on all pre-treatment covariates but one. This holds for the sample consisting of participants included in the experimental design before attrition and the sample used in analysis after attrition to estimate treatment effects. Neither sample was confounded by these pre-treatment variables. The figure plots p-values from difference-in-means tests between the treatment and control group. The only pre-treatment covariate that had a significant difference between treatment and control groups was attitudes toward whether the transition between the Biden and Trump administrations would be smooth. That one covariate out of 46 tests has a significant p-value is not too alarming as this is something we would expect to occur in due to the nature of null hypothesis significance tests at the $\alpha = .05$ level.

Because all participants were assigned a partner, then those partnerships were assigned to converse or not, our pre-analysis plan stated that we would drop all partnerships (including those in the control condition) where one partner did not show up to the experimental portion of the study. This is how we end up with equal number of Republicans and Democrats in our sample. Also it is important to note that this attrition occurred pre-treatment, and thus is not influenced by seeing whether you were asked to talk politics or not.

Figure A.1: Balance in Partnership-Level Covariates Across Treatment Conditions
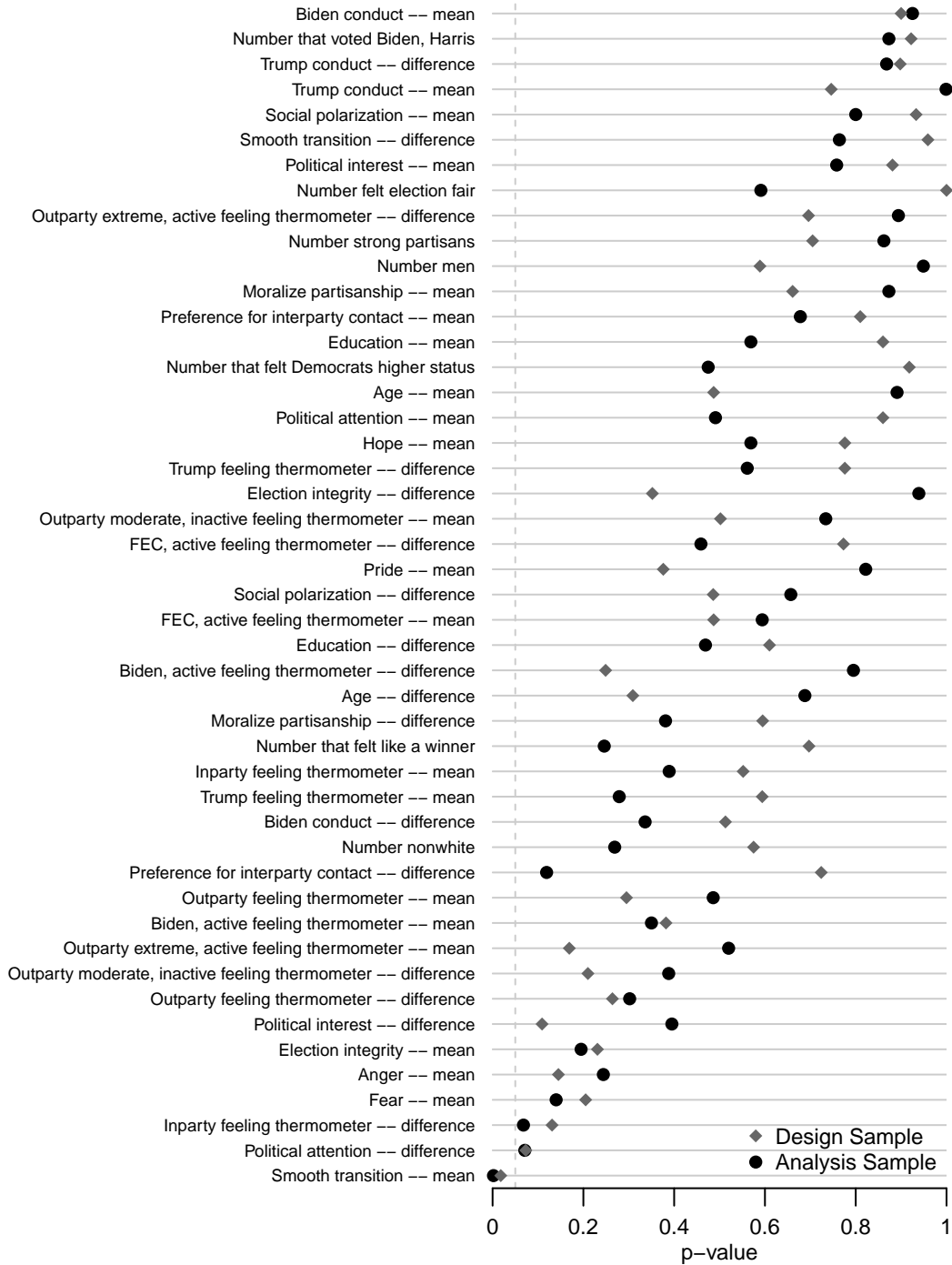
Note: Difference in means p-values assessing balance across treatment conditions for partnership-level pre-treatment covariates. Grey diamonds denote partners included in the design and black squares are partners in the sample in our analyses after attrition.

Appendix D   Experimental Conditions

Figure A.2 displays the instructions participants saw when they entered the chat app if they were assigned to the conversation condition. Figure A.3 shows the instructions those assigned to the short essay condition saw.
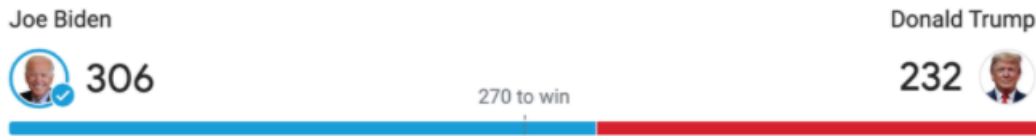
We chose a short-essay control for several reasons. The primary goal was to isolate the social experience of having a conversation. Therefore, we needed to create a control condition that primed many of the same features present in a conversation that are unrelated to the social element. Specifically, we wanted to create a control condition (1) where participants' partisan identities were primed as as they are in conversations about politics, and (2) where participants were asked to think about the same general topic (a salient, political topic) as in the conversation. We also wanted participants to engage in the activity for the same amount of time and interface with the same chat app technology. The short essay control accomplished each of these goals. The participants in the control condition had an identical prompt as in the conversation condition, but for two elements. First, the essay prompt introduces the out-partisan conversation partner, the first key ingredient to a social experience is having another person present. Second, the conversation prompt ask the participant have a conversation with that person, the second element needed to shift from an individual consideration of the 2020 election to a social setting where the 2020 election is discussed.

We chose not to use a control condition where people discuss the 2020 election but with a *same-party* member for several reasons. First and foremost, this kind of control condition puts participants in a social setting as well, and thus does not isolate the social experience of having a cross-partisan conversation, which was our goal. Second, the kinds of partnerships in this kind of control condition would be different from the kinds of partnerships in the treatment condition, thus the treatment and control groups would not be balanced on important characteristics, confounding results. For example, same-party partnerships might have similar personalities, similar issue attitudes, similar lifestyles, etc. whereas cross-partisan conversation would be more likely to differ on these pre-treatment traits. Therefore, using same-party conversation as a baseline for cross-party conversation does not allow us to isolate the role that *only* the partisan identity of the partner might play in how the conversation unfolds and it's effects on attitudes and behaviors because pre-treatment characteristics potentially correlate with treatment assignment and thus confound results. Comparing the effects of inter-party

to intra-party conversation is another interesting area for investigation, but was beyond the scope of our argument and design.

Figure A.2: Conversation Instructions (Treatment Condition)

After a long, hard-fought campaign, Joe Biden and Kamala Harris defeated Donald Trump and Mike Pence in the November 2020 election for president and vice presidentof the United States. While some battleground states were closely contested and legal challenges were raised in several states, the Electoral College formally voted on December 14, 2020 and Joe Biden won 306 Electoral College votes, clearing the 270 votes needed to win. After the Electoral College vote, top Republicans, including Senate Majority Leader Mitch McConnell congratulated Biden on his victory. Former Vice President Mike Pence declared Joe Biden the winner after Congress formally verified the Electoral College votes on January 6, 2021, despite challenges raised by some Republican legislators and a disruption in the process when rioters breached the Capitol building. Biden and Harris formally took office upon their inauguration on January 20, 2021.

Joe Biden   Donald Trump

306   270 to win   232

We've randomly assigned you to have a conversation with someone that belongs to or leans toward the **Republican** party. Please have a conversation about the result of the 2020 United States presidential election between Donald Trump and Joe Biden.

For example, you might talk about some of the following topics.

- How do you feel about the candidates' actions since the election?
- How do you feel about the state of the country these days—fearful, angry, hopeful, proud, or something else?
- Do you think the transition from the Trump Administration to the Biden Administration will go smoothly?
- Do you think that the election in November was administered fairly?

Please have your conversation by sending messages in the chat box below.

With bonus, this HIT is paying above minimum wage. We expect you to provide several comments and to utilize the full 8 minutes without large gaps of time. Participants who do so will receive a $1.00 bonus.

▶ Additional instructions about the chat app:

Write your reply...

Send

Done

Figure A.3: Short Essay Instructions (Control Condition)

After a long, hard-fought campaign, Joe Biden and Kamala Harris defeated Donald Trump and Mike Pence in the November 2020 election for president and vice presidentof the United States. While some battleground states were closely contested and legal challenges were raised in several states, the Electoral College formally voted on December 14, 2020 and Joe Biden won 306 Electoral College votes, clearing the 270 votes needed to win. After the Electoral College vote, top Republicans, including Senate Majority Leader Mitch McConnell congratulated Biden on his victory. Former Vice President Mike Pence declared Joe Biden the winner after Congress formally verified the Electoral College votes on January 6, 2021, despite challenges raised by some Republican legislators and a disruption in the process when rioters breached the Capitol building. Biden and Harris formally took office upon their inauguration on January 20, 2021.

Joe Biden   Donald Trump

306   270 to win   232

We'd like you to share your thoughts on the 2020 United States presidential election between Donald Trump and Joe Biden in an individual short essay.

For example, we'd like you to write about some of the following topics.

- How do you feel about the candidates' actions since the election?
- How do you feel about the state of the country these days—fearful, angry, hopeful, proud, or something else?
- Do you think the transition from the Trump Administration to the Biden Administration will go smoothly?
- Do you think that the election in November was administered fairly?

Please answer these questions by sending messages in the chat box below.

With bonus, this HIT is paying above minimum wage. We expect you to provide several comments and to utilize the full 8 minutes without large gaps of time. Participants who do so will receive a $1.00 bonus.
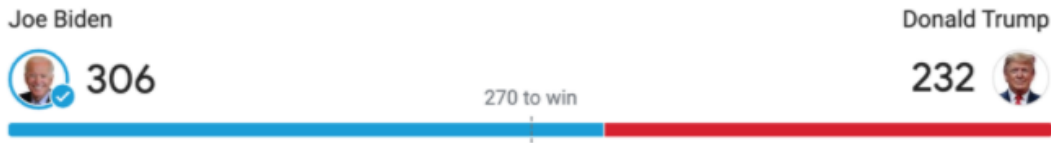
▶ Additional instructions about the chat app:

Write your reply...

|

Send

Done

Appendix E Example Conversations and Short Essays

Below are examples of conversations and short essays in response to the prompts shown in Figure
A.2 and Figure A.3. The length of the conversation and engagement between users is representative
of the sample of conversations. Moreover, the participants in the sample stay on topic, demonstrated
in these examples.

Conversation 1

user1	Hello
user2	hello
user2	I think Trump egging on that the election was stolen was very irresponsible
user2	because QANON followers really ate that up
user1	Rather unique election it was.
user1	Yeah, completely agree on that.
user1	News nowadays has become a cesspool.
user2	I dont really think it was stolen even though I am republican
user1	I believe there was probably fraud (just like in any election ever) but not enough to turn the outcome.
user2	I think it was done fairly but also from hearing everyone saying there was fraud I cant tell
user1	Yeah. I believe turnout was a bit higher too though, because of mail-in voting. More people were willing to actually participate because of lockdowns, etc.
user2	yeah so its kind of hard to tell
user2	I also think the state of the country right now is pretty bad even with Biden
user2	it looks pretty bleak
user1	If it were a Covid-free, usual year, I would say Trump may have likely won a second term.
user2	I 100 percent agree
user1	For sure–Biden or not, the country has a ways to go to recover.
user2	Has the transition been smooth
user1	I lean democrat, but I'm registered independent.
user2	hasnt biden overturned a lot of stuff or something
user1	Yeah, his first week was overturning quite a bit.
user1	Too much

Conversation 2

user1	Hi
user2	Hi
user1	Welcome, how are you today?
user2	I'm good. How did you feel about the outcome of the presidential election?
user1	I am okay with the result. I was not a fan of Trump
user2	I'm not a fan of Biden. I was for Trump. How do you feel Biden has done since taking office? Are you satisfied with his actions so far?
user1	I am a moderate so I like the things he has done with the vaccine priority. I worry about the cost of the relief package but somethings need to be done. How about you
user2	I'm worred about the cost of the relief package too, but it is definitely way overdue. There are a lot of people worried about having enough food and even rent to be able to have a place to live.
user1	agree
user2	So far as the vaccine, I think Biden is trying to take credit for a lot of things Trump put into place with it.
user1	Trump deserves a lot of credit for "Warp Speed" but I don't think the implementation plan was very good
user2	No, but Biden hasn't done well with an implementation plan either.
user2	And the restrictions in some states are killing the economy.
user1	I agree. It seems like things are getting much better though. My state opened up appts for my age group today and I got an appointment easily only for Wednesday
user2	I hope so. I'm glad you got an appointment easy, so many people haven't been able to.
user1	It has been nice chatting with you

Conversation 3

user1   Hello anyone there?

user1   It seems we have to chat about the recent election of President Trump vs. Biden.

user1   Do you have any thoughts about the senility of Biden? Does it look to you like he is being handled like anold puppet?

user2   I do believe that Biden is a little senile and definitely not completely fit to be president but its way betterthan dealing with Trump and his insanity.

user1   You forgot the cretin and corrupt parts about Joe.

user1   Even Obama called him stupid and Harris called him racist.

user1   I will be surprised if old Joe finishes his term. He's been brain damaged by strokes too.

user1   The Chinese Coronavirus might even finish old Joe off as well.

user2   Trump is also extremely corrupt and only in politics for power and attention like all politicians, Im not saying I support Biden and I do agree that hes definitely a racist but yet again its way better than dealing withTrumps policies and unstable nature. Biden sucks, I didn

user2   I didnt vote for Biden, I voted against Trump.

user1   There is likely fraud in the election and the excuse of a Chinese virus that has in a nation of over 333 millionpeople but killed less than 500,000 was a big help to perpetuate the fraud.

user1   Sounds like a lot of people in how you voted and I am not surprised.

user1   I will not be surprised if the election will be called tainted in the annals of history.

user2   Thats just more Trump brainwashing and nonsense. The virus is way worse here because of Trump and his awful handling of it and the amount of deaths due to this pandemic is astronomical for a disease with the advancements of modern medicine. The poor people who are stupid enough to think its a conspiracy or was created by China need to be educated about how to better check their news sources. The election wasnt tainted, thats just stupid.

user1   It was wrong

user2   It definitely wasnt, there is no evidence of actual voter fraud and every single case has been disproven orthrown out of court.

user1   They will be proven later no doubt but for now oh well.

Conversation 4

user1   Hello?
user2   Hello, how are you?
user1   I am doing well, how are you?
user2   Doing well.
user1   How do you feel about the state of the country these days?
user2   Going in the wrong direction. How about you?
user1   I am worried about the seeming lack of ability of people in power with differing viewpoints to work politelytogether to get jobs done.
user2   I agree with on that. I think the focus right now should be the people. A lot of people are hurting right nowand they are wasting time.
user1   Do you thing the election in November was fair?
user2   I do think it was fair. And you?
user1   I think that there is no evidence that it was not run fairly, or it would have come to light by now.
user2   Exactly. To much time was spent on election results and not enough on what was needed in this country.
user1   Thanks for the polite discussion, looks like time is almost up.
user2   Good chatting with you.

Short Essay 1

I find this topic to be very interesting. Do I think that the transition from the Trump to Biden Administration will go smoothly? No, I do not. I think this is because of the massive difference in changes between Biden and Trump. Some of the things Biden called for right off the bat were rather shocking to some. I think anytime you go from a Republican to Democrat and vs, it is not going to be entirely smooth.

The way I feel about the candidates since the election is actually quite surprising to myself. Yes, I did support and vote Trump, but his reaction to the election results was sickening. I fully think he acted like an immature monster, and not going to Biden's inauguration speaks volumes on his character. I think Biden accepted and began his presidency with grace, and that is something I greatly appreciate.

I am in fear for the state of this country these days, but that has been going on for a year now. I think Covid- 19 has had a detrimental effect for so many people, and my particular state (MI) has suffered. I am hopeful our country will get back to normal, but fearful for some of the things Biden wants done. However, I am not doubting him entirely, I would like to see what happens.

Yes, I think the election in November was mostly administered fairly. My biggest concern is the mail-in voting. It simply makes more sense that voting in person is the safest, most effective way to vote. I do fear with mail in voting that some people voted twice, or a deceased person's information was used. However, given the pandemic, it was necessary to offer especially for our most vulnerable.

Short Essay 2

I definitely think that the election was administered fairly. He is just a sore loser. I feel many things when i think about the state of this country. I feel fearful because of where it was heading but I'm also fearful about the future. I am moderate/middle of the raod so I agree with some things about each party but disagree on other things. I lean closer to the Republican side though.

The transition is seeming to go smoothly as we haven't heard much from Trump since the election

I think Biden is doing well so far, but only time will tell

The only negative I see Biden curently dealing with is that he can't promise an exact amount of vaccines given by this date. I don't think it's possible for him to know tha

I don't exactly feel proud of this country because of the last few months of recent events. How an I be proud of something that's embarrassing? Other countries sometimes look to use and if that's all they see, the riots and such, then that is an awful, awful representation of the United States

Short Essay 3

I still believe that the November 2020 election was a huge scam. There were so many instances of fraud that it blows my mind that the media was able to sweep it all under the rug. From hiding ballot counting from those assigned to the role, thousands of votes coming in in the middle of the night when the democrats realized they were behind, to "glitches" in voting machine software, the entire election should have been thrown out until weget a reliable system. I am still in awe of the news that came out that ANOTHER COUNTRY interfered with our election, and that information was never covered by the mainstream media. Americans today are so worried about their online presence that they won't tear their eyes away from their screens and see what is actually happening around them. They will not wake up until all of our freedoms as Americans are slowly stripped away and we become like every other country where you cannot defend yourself or your beliefs. By then it will be too late.

Short Essay 4

I think it was a very fair election and was very happy that Biden won. I feel that the country is headed in the right direction and we can gain our place back in international affairs. The transition went smoothly, although it was marred by the Capital protests. I am very happy Trump is largely off the gird, and Biden is an "A4" present, meaning, he is not constantly in the headlines. I am also happy that politics and government is not a constant issue in the news, as the major media outlets are moving more to hopeful news about the vaccine trials, international issues, other domestic news and entertainment/sports stories. It was a very bad 4 years under Trump, and I can only see our country moving in one direction and that's up.....On a personal level, all of my older relatives, and most of my older friends have had the COVID vaccine, and I personally do not know anyone who has even been diagnosed with the virus. Under Biden's direction, I think vaccine distribution will only get better and better..

Appendix F Robustness to Different Operationalizations of Feeling Thermometers

We also examined the extent to which our results were robust to different operationalizations of outparty affect. Consistent with recent work (Druckman et al. 2022), we find that outparty affect was slightly warmer for ideologically moderate, politically inactive outpartisans, compared to ideologically extreme, politically active outpartisans, at baseline. However, we find that the average treatment effect of conversation holds across all operationalizations of outparty affect. Table A.2 shows that regardless of whether individuals were thinking about (1) outpartisans across the country, (2) ideologically extreme, politically active outpartisans, or (3) ideologically moderate, politically inactive outpartisans, we find that conversations with outpartisans can increase positive feelings toward each of these characterizations.

Table A.2: Treatment Effects Using Different Operationalizations of Outparty Affect

| | *Outparty affect considering...* | | | | | |
	Outpartisans (main result)		Moderate, inactive outpartisans		Extreme, active outpartisans	
Conversation	6.22*	7.51*	6.68*	7.04*	4.13*	4.46*
	(1.41)	(1.82)	(1.41)	(2.05)	(1.35)	(1.76)
Partisanship (Rep.)		0.31		1.98		1.42
		(1.35)		(1.77)		(1.49)
Conversation x Partisanship (Rep.)		−2.58		−0.71		−0.67
		(2.11)		(2.75)		(2.35)
N	578	578	578	578	578	578

Note: * $p < 0.05$. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition. Conversation is 1 if the participant engaged in conversation with outparty member and 0 if the participant wrote an individual short essay. Partisanship is 1 if Republican and 0 if Democrat.

Appendix G Robustness to Affective Polarization As Difference Between Inparty and Outparty Affect

While we pre-registered using the outparty feeling thermometer to measure affective polarization as levels of outparty animosity, one could also use the difference between inparty and outparty feeling thermometers to measure affective polarization. To ensure our approach is robust to both measurement strategies, we report average treatment effects and heterogeneous treatment effects by partisanship in Table A.3 using the pre- and post-treatment change in affective polarization, where affective polarization is measured as the difference between inparty and outparty affect. The negative treatment effect estimates indicate that conversation caused the gap between inparty and outparty evaluations to shrink, relative to the control condition. We find that the effect of conversation on decreasing Democrats' affective polarization was slightly stronger, although not statistically significant. These results are thus consistent with our findings that conversation decreased outparty affect, although not to different degrees conditional on partisanship. We note that this robustness check was not pre-registered.

Table A.3: Results Robust to Affective Polarization Measured as Difference Between Inparty and Outparty Affect

	Affective Polarization	
Conversation	-5.82^*	-7.98^*
	(1.46)	(1.98)
Partisanship (Rep.)		-0.07
		(1.47)
Conversation x Partisanship (Rep.)		4.32
		(2.51)
Num.Obs.	410	410

Note: * $p < 0.05$. Models assess treatment effects for change in affective polarization, where affective polarization is measured taking the difference between inparty and outparty feeling thermometer ratings. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Appendix H   Interpreting the Interaction Term

We designed our study to be powered to detect an interaction of size -4.00 based on results from similar studies. Please see our pre-registration for the simulation study we ran to determine the sample size needed for 80% power for this coefficient size. Importantly, we determined our sample size for our analysis would need to be 584 participants, balanced across conditions. Our analysis sample includes 578 participants, with slightly more in the treatment condition than the control, giving us confidence our study was well-powered.

We determined the coefficient size of -4.00 *a priori* through discussion of what would be a substantively interesting difference between partisans based on coefficient sizes from previous studies using feeling thermometer measures. Our discussion of what would be a substantively meaningful difference between partisans also considered that our treatment was particularly intense. We expected partisan competition to be at extremely high levels immediately following Joe Biden taking office and the events of January 2020. Moreover, the topic of conversation was the electoral competition itself (the 2020 election). Despite the intense treatment, our interaction effect was only -2.58 shown in Table A.14 for outparty affect, our main outcome of interest. We were not powered to detect a coefficient this small. Therefore, we encourage future studies to build on this result and consider conditions that may lead to impactful differences in the effectiveness of cross-partisan discussion.

Appendix I   Measuring Partisan Group Threat

In this paper, we examined whether cross-party conversation reduced affective polarization differentially between Democrats and Republicans. We argued that in contentious political environments, such as the wake of the 2020 election and January 6th insurrection, partisan identity and group threat might be especially salient. Given the well-documented effects of partisan group threat on political behavior (Huddy 2013; Huddy, Mason and Aarøe 2015; Long, Eveland Jr and Slater 2019; Mason 2016), we expected that when partisans experience different levels of threat and reassurance, they will have unique reactions to cross-partisan conversation. Specifically, threatened groups – such as those who recently lost salient elections – will be less likely to reduce their outgroup animosity than reassured groups – such as those who recently won elections.

Measuring emotions, such as feeling threat, is no easy task. In this section, we provide additional justification for our design choice to examine differential feelings of group threat using partisanship. We rely heavily on previous research to support our conceptual approach and the measures we use to validate partisan group threat in our sample.

Conceptually, most work points to social identity theory as the framework through which to understand partisan group threat. For example, Huddy and Bankert (2017) write:

> The motivational underpinnings of social identity theory are *central* to understanding its expressive nature and ability to motivate political action...The maintenance of positive group distinctiveness is an active process, especially when a group's position or status is threatened, helping to account for the dynamic nature of partisan political activity (Huddy 2013; Mackie 2000). Elections pose threats to both a party's power and less tangible goods such as group members' collective social standing...

This line of research suggests that many features of the political environment can influence partisan group threat. For example, Mason (2016, p. 4) explains: ". . . These status threats can be induced by nearly anything that suggests that one party will emerge victorious. Elections, debates, elite policy disputes, or simple media rhetoric—all act as threats or reassurances to the standing of one party or the other." Political status of one's ingroup (their party) is thus one key factor perceived as at risk of weakening, relative to other groups who might gain dominance.

Previous studies have captured this notion of partisan group threat in a number of ways. For example, Long, Eveland Jr and Slater (2019) operationalize partisan group threat as being in a political minority within one's immediate social network and within their broader geographic context (county). Other work has focused on measuring emotions known to correlate with threat, such as anger, or reassurance, such as pride. Mason (2016) uses ANES data to examine emotions toward outparty and inparty presidential candidates, suggesting that when respondents reported feeling angry toward the outparty candidate and pride toward the inparty candidate, this could be an admittedly crude measure of threat.

Most relevant for our work are studies that have experimentally induced feelings of threat. In addition to the conceptualization of partisan group threat discussed above, we can examine how these studies design their vignettes to elicit feelings of threat. These studies chose to manipulate threat by highlighting one party's decline and another's rise. For example, Mason (2016) used the following experimental stimuli to induce partisan group threat for Democrats and Republicans:

The following statement recently appeared on a Democratic blog: "2012 is going to be a great election for Democrats. Obama will easily win reelection against whatever lunatic the Republicans run, we are raising more money than Republicans, our Congressional candidates are in safer seats, and Republicans have obviously lost Americans' trust. Our current Congress is proving to Americans that Republicans do not deserve to be in the majority, and Americans will make sure they're gone in 2012. Finally, we'll take the Congress back and won't have to worry about the Republicans shutting down government anymore! I'm glad that Americans have finally returned to their senses. Republicans should get used to being the minority for the foreseeable future. Democrats will hold our central place in the leadership of the country. Obama 2012!!"

The following statement recently appeared on a Republican blog: "2012 is going to be a great election for Republicans. We're going to defeat the hardcore socialist Obama, we are raising more money than Democrats, our Congressional candidates are in safer seats, and Democrats have obviously lost Americans' trust. Our current Congress is proving to Americans that Democrats do not deserve to be in the majority, and Americans will

make sure they're gone in 2012. Finally, we'll take the government back, and we won't have to worry about Democrats blocking us at every turn! I am so glad that Americans have finally returned to their senses. Democrats should not get used to running the government. Republicans will take back our central place in the leadership of the country. Defeat Obama in 2012!!"

This stimulus is similar to that used by Huddy, Mason and Aarøe (2015) and other research, which has examined group threats in other contexts, such as racial group threat (Pérez 2015). The key is that these studies design an exogenous prime of threat that describes a scenario in which one's ingroup party is losing or prime reassurance by describing their ingroup party as winning.

Following very closely with the conceptual framework used by previous research, we argue in this paper that following an election, partisans should naturally experience feelings of threat when their party lost and reassurance when their party won. We saw value in studying the effects of organic feelings of threat spurred by the political environment rather than randomly assigning a threat stimulus as prior work as done. However, because we did not control the threat stimuli like prior work, we carefully included three "manipulation check" questions that tap into the key components of threat, as used by previous research described above. We call these "manipulation check" questions because we expected the environment to prompt differential feelings of partisan group threat, so we asked these questions in order to assess whether our partisan participants indeed felt as we expected. These are only three such questions one might ask about threats and reassurances from the election and inauguration, and other questions could have been used.

1. Winning and losing

Our first question directly measured whether people felt more like "winners" or more like "losers" following the election, which is explicitly tied to the importance of winning and losing in previous research on threat. Specifically, we asked respondents: "Politics is often compared to sports. After the 2020 presidential election, did you feel more like a "winner" or more like a "loser" in politics?" Note that this question directly asks about their response to the election, meaning that it should not capture "winning" or "losing" feelings in response to something else. Moreover, our choice to specifically reference sports comes from previous work which ties threat and reassurance within a

political context to sports. For example, when developing their theory of partisan identity strength and threat and reassurance, Huddy, Mason, and Aarøe (2015, p. 3) write: "But strong partisans should also respond actively to reassurance because such expressive behavior feels inherently good in much the same way that assured victory increases fan enthusiasm and turnout at a sporting event."

Table A.4 shows the results. We find that 93% of Democrats reported feeling more like a winner after the election, while only 15% of Republicans reported feeling this way.

Table A.4: Feeling like a "Winner" or "Loser" by Partisanship

	Democrats	Republicans
More like a "winner"	.93	.15
More like a "loser"	.07	.85

2. Changing Status

Our second question was a bit more abstract for participants to answer, but tied directly to the theoretical construct of group threat, which emphasizes how threat is spurred by expecting or experiencing the changing status of groups within society. Specifically, we asked respondents: "Which party do you think has higher political status right now? By higher status, we mean more political power, more political influence, or more of an advantage in politics." This question explicitly highlights that one group has more influence than another, which follows directly from research on group threat. If partisans perceive the outgroup as having more political influence, they should experience feelings of threat; whereas if they perceive the ingroup as having more political influence, they should experience feelings of reassurance.

We find that both Democrats and Republicans perceived Democrats to have higher political status in January 2021. Specifically, 92% of Democrats and 92% of Republicans reported that Democrats had higher political status, as shown in Table A.5. This overwhelming consensus further suggests that partisans in our sample perceived the political environment to favor Democrats over Republicans, which could in turn cause Democrats to feel reassured and Republicans to feel threatened.

Table A.5: Party Perceived to have Higher Status by Partisanship

	Democrats	Republicans
Democrats have higher political status	.92	.92
Republicans have higher political status	.08	.08

3. Feeling threat

Our third question directly asked participants about their feelings of threat after the 2020 election. Specifically, we asked: "When Democrat Joe Biden won the election and became President of the United States, how did you feel? Check all that apply. [Threatened, Anxious, Angry, Afraid, Disgusted, Reassured, Happy, Relieved, Hopeful, Enthusiastic]"

Looking specifically at the proportion of Republicans and Democrats who reported that the felt threatened, we found that 0.6% of Democrats felt threatened and 34% of Republicans felt threatened, a difference in proportions that is statistically significant. Although less than half of Republicans in our sample reported directly feeling threatened, it is noteworthy that functionally 0 Democrats reported feeling threatened. This is especially important in the context of the 2020 election when Republicans were challenging the election results, which could have been viewed as a threat to Democrats. This does not appear to be the case.

Similar patterns hold when directly asking participants about feeling reassured after the 2020 election. We find that 49% of Democrats report feeling reassured, while only 4% of Republicans felt this way, a difference in proportions that is statistically significant.

Moreover, previous research on threat and reassurance has used other emotions related to the constructs. For example, Mason (2016) uses ANES data to measure threat and reassurance. She measures threat to a party's status by using a respondent's self-report of whether they felt angry toward their outgroup presidential candidate; and measures reassurance with the respondent's self-report of whether they felt proud of their ingroup presidential candidate (p. 7). She notes in Footnote 3 (page 7) that she also examined "hopeful" and found similar results as "pride." We asked respondents about several different emotions that could tap into threat and reassurance. Table A.6 shows the proportion of Democrat and Republican respondents who reported feeling each emotion we asked about, sorted into threatening and reassuring emotions. For each threatening emotion, a significantly greater proportion of Republicans reported feeling it; whereas for each reassuring

emotion, a significantly greater proportion of Democrats reported feeling it.

Table A.6: Threat and Reassurance Emotions by Partisanship

	Democrats	Republicans
Threatening emotions		
Threatened	.006	.34
Anxious	.09	.47
Angry	.006	.31
Afraid	.02	.39
Disgusted	.02	.43
Reassuring emotions		
Reassured	.49	.04
Enthusiastic	.37	.03
Happy	.54	.05
Relieved	.80	.10
Hopeful	.73	.19
At least one threatening emotion	.11	.80
At least one reassuring emotion	.97	.26

Vote Choice

Last, we turn away from our three "manipulation check" questions that stem from our conceptualization of partisan group threat to assess who in our sample was potentially feeling threat by examining self-reported vote choice. The conceptual framework laid out above suggests that supporting the losing candidate is an indicator of experiencing threat and supporting the winning candidate would be an indicator of experiencing electoral assurances. Therefore, it could be the case that Republicans and Democrats in our sample were not experiencing threats and reassurances as we expect because they did not vote along party lines in the 2020 Presidential election. However, we show in Table A.7 that 92% of Democrats and 76% of Republicans voted in line with their partisanship. This finding of who may be experiencing electoral threats and reassurances, coupled with our three manipulation check questions of what "threat" should look like in our context, provides robust evidence for using partisanship to assess heterogeneous treatment effects of conversation based on partisan group threat.

In sum, our conceptual framework for partisan group threat builds directly on previous research, we use several measures stemming from this framework to assess whether partisanship is a good way

Table A.7: Vote Choice by Partisanship

	Democrats	Republicans
Biden, Harris	.92	.13
Trump, Pence	.01	.76
Other	.02	.03
Did not vote	.05	.08

to capture differential feelings of threat at the time of our study, and we show that partisanship is highly correlated with each of these measures. Moreover, because we have these three theoretically-driven operationalizations of threat, we assess heterogeneous treatment effects of conversation using each of these measures in place of partisanship in Appendix J. In each case, we fail to find feelings of threat condition the effect of conversation on decreasing affective polarization, consistent with the results reported in the paper using partisanship.

Appendix J   Robustness to Different Operationalization of Partisan Group Threat

In this Appendix we show our results are robust to partisan group threat operationalized in ways other than partisanship. We replicate our results when operationalizing partisan group threat via our three variables, which we pre-registered as "manipulation checks," and an additional robustness check measure, which was also pre-registered: (1) participants' personal feelings of being a winner or loser in politics after the 2020 election, (2) participants' perceptions of the parties' political status at the time of the survey, (3) feelings of threat after the 2020 election, and (4) self-reported vote choice. Please see Appendix I for a deeper discussion of each of these measures as well as summary statistics for how these measures correlate with partisanship.

We focus on Hypotheses 4-6 where we change the operationalization of threat as the moderating variable and evaluate whether heterogeneous treatment effects of conversation are robust.[1]

1. Winning and Losing

First, we replicate our main results, but rather than operationalize threat by assuming Republicans experienced threat and Democrats experienced reassurance, we operationalize threat as anyone who indicated they felt like a "loser" after the 2020 election (*Felt like a "Loser"* = 1) and reassurance as those who felt like a "winner" (*Felt like a "Loser"* = 0).

Table A.8 reports results of this operationalization of threat in the first four columns. We fail to find support for our pre-registered hypotheses that threat conditions the effect of conversation on outparty affect, social polarization, and perceptions of election integrity.

We do find that support for Democratic values now has significant heterogeneous treatment effect estimates. Threat leads to a significantly smaller effect of conversation on support for Democratic norms than does reassurance. This may have been due to a more acute reminder the conversation provided about the threats to American democracy surrounding the 2020 election. However, recall this was an exploratory analysis, and this section shows that heterogeneous treatment effects are *not* a robust finding to other operationalizations of partisan group threat. We caution readers from

---

[1]These robustness check results also hold if we use partisanship as the moderator but limit our sample to only those partisans who experienced threat.

interpreting the results in Table A.8 as strong evidence that conversation differentially improves support for democratic norms for those feeling threat or reassurance, pending these results can be replicated in future work.

Table A.8: Results Operationalizing Threat via "Manipulation Check" Measures

	Outparty Affect	Social Polarization	Election Integrity	Democratic Values	Outparty Affect	Social Polarization	Election Integrity	Democratic Values	Outparty Affect	Social Polarization	Election Integrity	Democratic Values
Conversation	7.37*	0.08	−0.02	0.24*	6.96*	0.08	−0.006	0.19*	7.01*	0.10*	−0.007	0.19*
	(1.86)	(0.04)	(0.05)	(0.09)	(1.84)	(0.05)	(0.05)	(0.10)	(1.74)	(0.04)	(0.05)	(0.09)
Felt like a "loser"	−0.05	−0.05	−0.003	0.18*								
	(1.45)	(0.03)	(0.05)	(0.08)								
Conversation × Felt like a "Loser"	−2.46	0.01	0.07	−0.30*								
	(2.25)	(0.05)	(0.07)	(0.11)								
Outparty Higher Status					0.16	−0.06	0.01	0.16				
					(1.47)	(0.03)	(0.05)	(0.08)				
Conversation × Outparty Higher Status					−1.47	0.02	0.03	−0.17				
					(2.22)	(0.06)	(0.07)	(0.11)				
Felt Threat Emotions									−0.07	−0.06	0.004	0.10
									(1.54)	(0.04)	(0.05)	(0.09)
Conversation × Felt Threat Emotions									−1.97	−0.04	0.05	−0.23
									(2.36)	(0.06)	(0.07)	(0.12)
Num.Obs.	578	578	578	578	578	578	578	578	578	578	578	578

* $p < 0.05$

Note: * $p < 0.05$. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

2. Changing Status

Second, we operationalize threat using our "manipulation check" question asking about the relative status of the Democratic and Republican parties. We operationalized threat as anyone who reported their outparty had higher political status at the time. In other words, *Outparty Higher Status* equals 1 if a Republican that reported Democrats had higher political status or a Democrat that reported Republicans had higher political status at the time. *Inparty Lower Status* equals 0 if a Democrat that reported Democrats had higher political status or a Republican than reported Republicans had higher political status.

Results are shown in columns 5-8 of Table A.8. We fail to find this operationalization of threat conditions the effect of conversation on any of the four outcomes.

3. Threat Emotions

Third, we analyze participants who reported in the pre-treatment survey feeling threat and reassurance emotions as we would expect under our conceptualization of partisan group threat. Specifically, *Felt Threat Emotions* equals 1 if the proportion of threat emotions the respondent indicated feeling from Table A.6 was greater than the proportion of reassurance emotions. Table A.8 shows in the final four columns that we fail to find evidence of heterogeneous treatment effects of conversation due to threat operationalized in this way.

Vote Choice

Finally, Table A.9 reports that our results remain consistent when we probe the robustness of our results by looking for heterogeneity conditional on vote choice rather than partisanship. Specifically, we operationalize threat as anyone who reported voting for Trump and Pence. We operationalize reassurance as anyone who reported voting for Biden and Harris. We omit those who reported voting for other candidates or who reported they did not vote (a small percentage of the sample, as shown in Table A.7.) When operationalizing electoral threat in this way, we again find no evidence of heterogeneous treatment effects of conversation stemming from threat.

Table A.9: Results Operationalizing Threat via Vote Choice

	Outparty Affect		Social Polarization		Election Integrity		Democratic Values	
Conversation	5.81*	7.24*	0.09	0.13*	−0.05	−0.07	0.15	0.24*
	(1.92)	(2.37)	(0.05)	(0.06)	(0.04)	(0.06)	(0.09)	(0.12)
Trump Vote		0.34		−0.06		0.01		0.12
		(1.75)		(0.04)		(0.06)		(0.10)
Conversation × Trump Vote		−2.85		−0.06		0.05		−0.19
		(2.62)		(0.07)		(0.09)		(0.14)
Num.Obs.	400	400	400	400	400	400	400	400

Note: * $p < 0.05$. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Appendix K   Polarization Outcomes by Condition Across Time

To demonstrate that conversation is causing the observed treatment effects in Figure 1, we plot raw means and 95% confidence intervals for both polarization outcomes measured pre-treatment and post-treatment for the treatment and control conditions separately in Figure A.4. The first panel shows pre- and post-treatment means for outparty warmth for the conversation (treatment) group. The second panel shows pre- and post-treatment means for outparty warmth for the short essay (control) group. The third and fourth panels show the same, but for the social polarization outcome.

This figure shows little post-treatment change from pre-treatment levels of warmth and social polarization for the control group (panels two and four). This finding demonstrates that positive treatment effects of conversation are not driven due to the short essays *decreasing* outparty warmth, which perhaps could have happened if the control condition lead participants to reflect and become more upset toward outparty members. Instead, because there was little change pre-post for the control group, we can interpret our positive treatment effects as being driven by conversations *increasing* warmth toward the outparty.

Table A.10 further assesses how the conversation and short essay groups experienced each experimental stimulus. For each experimental group separately, we conduct a paired t-test to see whether the average *within-subject* change was significant, and whether it was positive or negative. The first row of Table A.10 shows that on average, participants in the conversation group increased their outparty warmth by 9.10 degrees—a statistically significant improvement. Likewise, participants in the short essay condition had a statistically significant increase in their outparty warmth, but to a much smaller degree of only 1.75 degrees on average. We see similar patterns for the social polarization outcome. The conversation led to a statistically significant improvement in participants' attitudes, while the short essay's effect is not distinguishable from 0.

This is evidence against the hypothesis that the control group *decreased* their outparty warmth. Therefore, we attribute the positive treatment effects of the conversation relative to the short essay to the conversation intervention increasing warmth.

Figure A.4: Means and 95% Confidence Intervals for Polarization Outcomes by Condition and Across Time
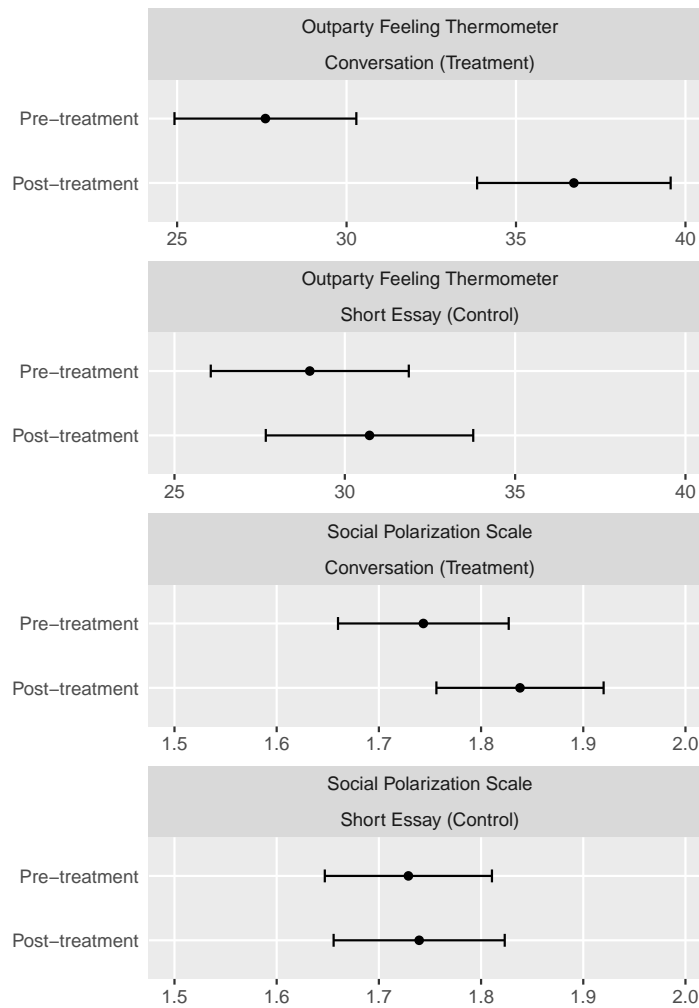
Table A.10: Paired t-tests for Polarization Outcomes by Condition

Outcome	Experimental Group	Est.	Std. Err.	t-stat	p-value	95% CI	N
Outparty feeling	Conversation	9.10	0.95	9.63	0.00	[7.24, 10.96]	294
thermometer	Short Essay	1.75	0.59	3.00	0.00	[0.60, 2.91]	284
Social	Conversation	0.09	0.02	4.16	0.00	[0.05, 0.14]	294
polarization scale	Short Essay	0.01	0.01	0.80	0.42	[-0.02, 0.04]	284

Note: Paired t-tests results examining within-subject change in the two polarization outcomes over time (before and after treatment).

Appendix L   Mechanisms

We examine several possible mechanisms, which we pre-registered as exploratory mechanism checks. Our experiment revealed that even in highly competitive contexts, conversation can increase outparty affect, regardless of partisanship and feelings of partisan group threat. There are many possible mechanisms through which conversation can work to reduce outgroup hostility, which we broadly group as the depth and quality of the connection from the conversation experience.

Depth of Connection

Previous research on intergroup contact suggests that contact is effective because it allows people to connect with one another despite their differences. The first dimension of connection we consider is the depth of this connection, and one way to consider the depth of connection is how much participants actually engaged in conversation with each other. The more individuals discussed, the more we should expect them to have opportunities to connect and subsequently reduce their animosity toward the outparty. We tested this by analyzing the transcripts of the conversations and short essays. The conversations occurred via an online chat, allowing us to easily measure the number of times each participant "spoke" in the conversation. In the control condition, we broke down the short essays into sentences.

In Table A.11, we first standardized the "turns" variable (sentences for the control group; speaking turns for the conversation group) to have a mean of 0 and a standard deviation of 1 for easier interpretation. In the first model, we show that the number of turns one takes in their short essay/conversation is not associated with outparty affect. Instead, when we interacted treatment assignment and the standardized number of turns taken by each participant, we found that in the treatment group, each an increase in one standard deviation of the turn taken was associated with a 5.5 unit increase in outparty affect. But, there was no statistically significant association between the number of sentences written and outparty affect in the control group. In all, it was only in the interactive, conversation group where more engagement was associated with improved outparty affect.

Moreover, we did not find that Republicans and Democrats participated at different rates. In coversation settings, the mean number of turns taken for Democrats was 9.30 and for Republicans

Table A.11: Length of Conversation and Outparty Affect

	Outparty Affect	
Number of Turns	−0.12	−0.89
	(1.11)	(1.18)
Conversation		6.53*
		(1.67)
Number of Turns x Conversation		5.54*
		(2.13)
Num.Obs.	522	522

Note: * p < 0.05. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

was 9.46, with a t-test p-value of .78.

The second way we consider depth of connection is whether individuals shared personal information about themselves (self-disclosure), see things from their partner's point of view (empathy), and learn something new about the outgroup (learning). We measured each of these variables on a 5-point likert scale.[2] These questions were asked post-treatment and only of participants in the conversation group. We are very cautious in interpreting the results from this analysis. We do not intend them to be interpreted causally in any way; but rather, we interpret them as correlations that are suggestive as possible mechanisms future research could explore.

Figure A.5 shows the coefficients from simple bivariate linear regressions where the independent variable is the mechanism of interest and the dependent variable is the change in outparty affect. For each mechanism, self-disclosure, empathy, and learning, we observe a positive, statistically significant association with outparty affect. That is, the more someone shared personal information with their partner, felt that they could see things from their partner's point of view, and learned something about the outparty, the more they also improved their attitudes toward the outparty.

As with our experimental findings, we did not find any evidence to suggest that partisanship conditioned reports of connection and improved outparty attitudes. We interacted each mechanism with partisanship, as we were interested in whether Democrats who reported learning, empathizing,

---

[2]We measured self-disclosure using a question adapted from Laurenceau, Barrett, and Pietromonaco (1998): "How personal was the information you disclosed to your partner?" (1=not personal at all; 5=extremely personal). We measured empathy using a question adapted from Davis (1983): "How easy was it for you to see things from the point of view of your conversation partner?" (1=extremely difficult, 5=extremely easy). We measured learning by asking respondents to report how strongly they agreed (1=strongly disagree, 5=strongly agree) with the statement "I learned something about [Democrats/Republicans] from the conversation."

Figure A.5: Exploratory Mechanisms for Improved Outparty Affect
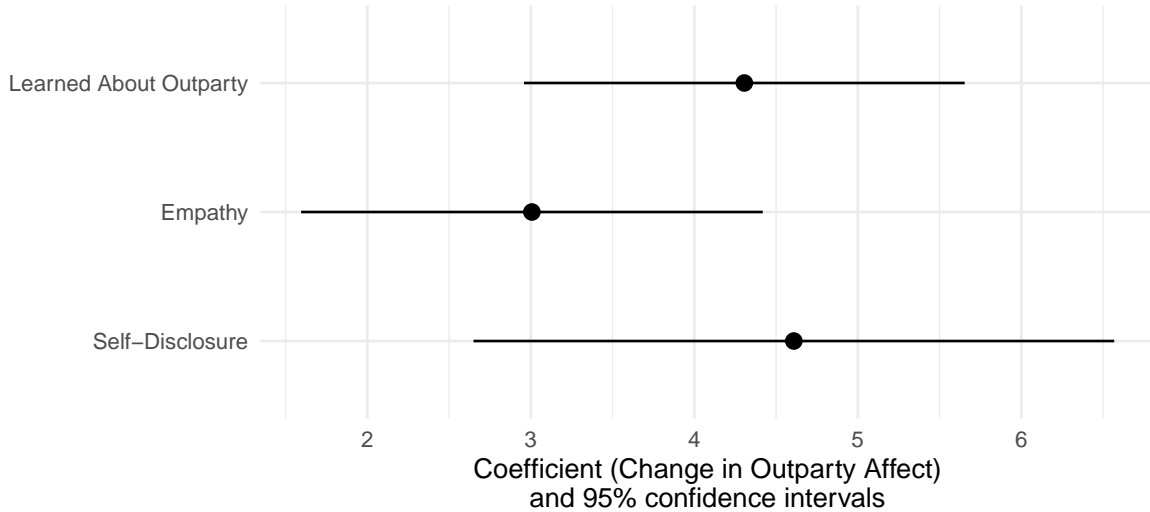
or self-disclosing improved outparty affect more than Republicans who reported these conversation experiences. We find no evidence that self-reporting these dimensions of connection had differential effects for Democrats and Republicans; however, we again interpret these results as exploratory and warranting future research.

Finally, we also examined if depth relative to one's partner helped explain changes in outparty affect. We create an indicator for whether a participant participated more, in terms of number of turns, whether in the treatment or control condition. In the first model in Table A.12 we show that talking more than one's partner is not associated with more change in outparty affect. Likewise, in the second model, we interact the indicator for talking more with an indicator for treatment assignment. We find no evidence that talking more than one's partner has an effect in either the treatment or the control condition.

Quality of Connection

Closely related to the depth of connection between the participants is simply the manner in which they connected—was it a positive or negative experience? Recent evidence suggests that the anticipated tone of a conversation, as civil or heated, affects individuals' willingness to engage in a political discussion even more than simple disagreement (Connors and Howell 2022). Moreover, previous research on intergroup contact suggests that its effectiveness is conditioned on whether the

Table A.12: Talking More than Partner and Outparty Affect

	Outparty Affect	
Talked More than Partner	1.01	−0.30
	(1.15)	(1.49)
Conversation		4.61*
		(2.00)
Talked More than Partner x Conversation		2.53
		(2.26)
Num.Obs.	522	522

Note: * p < 0.05. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

experience is civil and positively valenced. To test this, we return to both the self-reported data on the discussion experience and the transcripts.

First, we asked respondents to simply report whether they had a positive or negative experience in the conversation, following a method used by (Barlow et al. 2012), and expecting that those with positive contact experiences would have greater improvement in outparty attitudes. Respondents' blunt assessments of whether they had a positive or negative experience was strongly associated with change in outparty affect. The vast majority of respondents reported that they had a positive experience with their partner. Those who reported having a positive experience improved their outparty affect by 10.7 points, compared to those who had a negative experience, who decreased their outparty affect by 2.4 points. The change in affect among those who had a negative experience is not statistically distinguishable from zero. This suggests that negative conversation experiences are not associated with polarized attitudes, while positive experiences are associated with improvements in attitudes. We again caution readers against interpreting these results causally.

When we analyze the free-response descriptions of how participants felt over the course of the conversation, we find - once again - that the positive experiences were associated with more positive outparty affect. We used a sentiment dictionary to estimate the positive and negative sentiment of these descriptions (Young and Soroka 2012). We found that 37% of the descriptions were more positive than negative when describing how they felt at the *start* of the conversation, but when describing how they felt at the *end* of a conversation, 67% of the responses were more positive than negative. Here too, the more positive the change in feelings, the more people improved their outparty affect. These results are correlational, so we encourage future research to experimentally

investigate these concepts as potential mechanisms of politically-charged interparty contact.

With a bit more precision, we return to the full transcripts of the conversations and short essays. We estimated the sentiment of the transcripts at the room-level by aggregating the positive sentiment words of the two participants. In contrast to the self-reported experiences, we find no evidence that the positive or negative tone of the conversation or essay was associated with outparty affect. Once the number of turns - which we analyzed as part of the "depth" of connection - was included in the models, the relationship between sentiment and outparty affect disappeared. This suggests that the civility of the language used is perhaps less important than the amount of language used in shaping outparty affect.

Appendix M   Durability

70.9% of participants returned for our follow-up survey at least three days after the experiment. Table A.13 reports the details for our findings that treatment effects for the main outcome of interest—change in outparty affect—are durable for at least three days after treatment. We did not ask about our other two main outcomes of interest (social polarization and perceptions of election integrity). We find a positive and significant treatment effect persists (Cohen's $d = .24$), and we again fail to find any evidence of a heterogeneous treatment effect across partisans.

Table A.13: Treatment Effects of Conversation Persist for At Least Three Days

	Outparty Affect	
Conversation	4.54*	5.31*
	(1.78)	(2.30)
Partisanship (Rep.)		2.57
		(1.83)
Conversation x Partisanship (Rep.)		−1.56
		(2.79)
Num.Obs.	410	410

Note: * p < 0.05. Models assess durability of treatment effects for main outparty affect feeling thermometer outcome. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Appendix N   Figure 1 Results

Figure 1 in the main text displays estimates of sample average treatment effects (SATEs), conditional average treatment effects by partisanship (CATEs), and heterogeneous treatment effects (the difference between the CATEs) for each of our three main outcomes of interest and one exploratory outcome of interest. Full tables of these results are in Table A.14 and Table A.15.

Table A.14: Sample Average Treatment Effects and Heterogeneous Treatment Effects by Partisanship

	Outparty Affect		Social Polarization		Election Integrity		Democratic Values	
Conversation	6.22*	7.51*	0.09*	0.09*	0.01	0.00	0.10	0.19*
	(1.41)	(1.82)	(0.03)	(0.04)	(0.03)	(0.05)	(0.07)	(0.09)
Partisanship (Rep.)		0.31		−0.05		0.02		0.08
		(1.35)		(0.03)		(0.04)		(0.08)
Conversation × Partisanship (Rep.)		−2.58		−0.02		0.01		−0.18
		(2.11)		(0.05)		(0.07)		(0.11)
Num.Obs.	578	578	578	578	578	578	578	578

Note: * p < 0.05. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Table A.15: Conditional Average Treatment Effects by Partisanship

	Outparty Affect		Social Polarization		Election Integrity		Democratic Values	
	Dem.	Rep.	Dem.	Rep.	Dem.	Rep.	Dem.	Rep.
Conversation	6.85*	5.59*	0.12*	0.05	0.03	−0.01	0.24*	−0.03
	(2.10)	(1.85)	(0.05)	(0.06)	(0.05)	(0.04)	(0.11)	(0.09)
Num.Obs.	289	289	289	289	289	289	289	289

Note: * p < 0.05. Conditional average treatment effects (CATEs) by partisanship. The model includes blocked fixed effects to reflect the design's randomization of treatment. Because these results are conditional on partisanship, there are no clustered standard errors (each participant's partner is an out-partisan and therefore not included in CATE estimation).

Appendix O Conversation and Short Essay Handcoding Procedure and Results

To aid participants in engaging with the topic of the 2020 election, we provided four questions to prompt discussion. The questions were presented in a random order across chatrooms so partners saw the same order, but different chatrooms saw the topics in a different order.

An undergraduate RA hand-coded each message sent in the conversations. To have a comparable unit of text to hand-code in the short essays, we parsed the short essay text at the sentence-level, and the RA hand-coded each of these units of text. The RA was given the four prompts the participants saw and was instructed to code a "1" if "the primary focus of the message was in response to this question" and 0 otherwise. In addition to the four prompts we provided to participants, we asked the RA to handcode for participants mentioning two events that were salient at the time: Joe Biden's Inauguration or the January 6th insurrection. The RA was also given the following instructions:

- Some messages might include content that could be in response to more than one of the four question prompts. You should enter a 1 for the topic that was more dominant in the message, perhaps being referenced first or in more detail than the other topic. All other prompt columns that were not the primary focus should be coded as 0. In other words, only code one question prompt for each message.

- Many messages will not reference any of the four question prompts. For example, there will likely be messages that are simply greetings (Hi!, Bye, etc.), other messages might indicate waiting for the other person to arrive (Hello? Anyone there?), and others could simply be entirely off-topic. All of these situations would simply get a 0 for all four of the prompt columns.

We drew the 2020 election prompts from questions asked in a Pew Research Center survey (Pew Research Center 2020). Appendix D shows the full instructions given to participants, and we list four prompts we provided participants here:

- **Transition to Biden Administration:** Do you think the transition from the Trump Administration to the Biden Administration will go smoothly?

- **Candidates' Actions:** How do you feel about the candidates' actions since the election?

- **Election Fair:** Do you think that they election in November was administered fairly?

- **State of the Country:** How do you feel about the state of the country these days–fearful, angry, hopeful, proud, or something else?

Table A.16 shows that the rates at which participants talked about these topics differs depending on whether they were assigned to a conversation or short essay. From reading the conversations and short essays, we know that one reason for this pattern is that conversations involve more "messages" devoted to greetings ("hi" and "bye") and social coordination ("do you want to go first?"). (See Appendix E for several examples.) This reduces the time and space used on the prompted questions relative to the short essays. That participants used some of their conversational time to engage in the social experience of having a conversation is a part of the bundled treatment of "conversation" we designed our study to test.

A second reason for this pattern is that we found that in the short essays, participants often crafted their response by responding to each prompted question in the order they were displayed to them. (Note that the order the prompted topics were displayed was randomized by chatroom.) Thus, Table A.16 shows that short essays discussed a higher number of prompted topics than did conversations.

Table A.16: Number of Topics Discussed by Experimental Condition

Number of Topics Discussed	Short Essay	Conversation
0	0%	2%
1	2%	30%
2	11%	43%
3	33%	18%
4	53%	7%
	100%	100%

Note: Table shows the rate at which chatrooms, either individual short essays or paired conversations, discussed any number of the four prompted topics.

However, raw counts of the number of topics discussed is only one measure of engagement with the exercise and the topic of the 2020 election. We also look at how many words and messages were sent across the short essays and conversations. We find that essays used 203 words on average while conversations used 304 words on average (or 152 per person). Coupled with the findings in Table

44

Figure A.6: Discussion of Different Topics by Experimental Condition
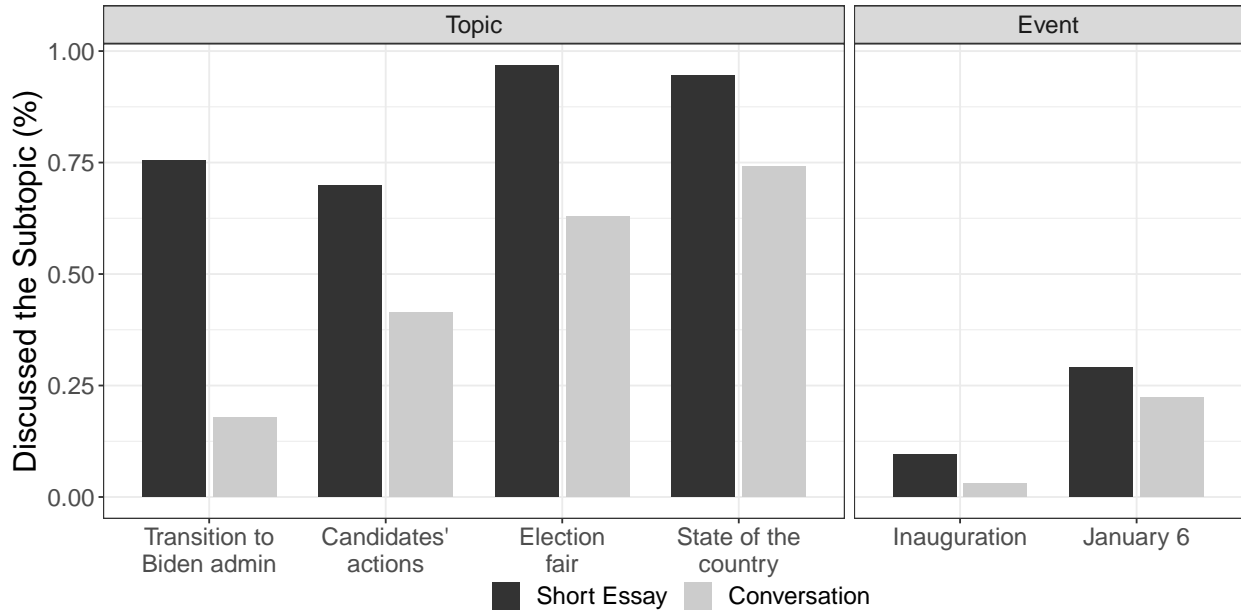
Note: Figure shows the proportion of chatrooms, either individual short essays (light gray bars) or paired conversations (dark gray bars), that discussed each of the four prompted subtopics and the two salient events at least once.

A.16 that show short essays discuss more of the prompted topics on average, we find that short essays do so to a lesser depth than do conversations which have more content exchanged.

We also explore differential use of the four topics in Figure A.6. As we show in Table A.16, a majority of the short essays (dark grey bars) touch on all four subtopics of the 2020 election, with nearly 100% of short essays discussing the fairness of the election administration and approximately 75% of short essays discussing the transition from the Trump to the Biden administration, for example. Conversations did not usually discuss all topics, with nearly 75% touching on the state of the country these days and only 18% discussing the transition to the Biden administration.

We did not explicitly ask participants to discuss the inauguration or January 6th, but we looked for reference to these specific events. Again, short essays brought these events up more on average than did conversations.

One might wonder whether conversation would have the positive effects we observe if participants conversed or wrote about the same subtopics regardless of if they were in a conversation or short essay. Therefore, we next test whether the subtopic(s) people chose to discuss or write about

helps to explain the positive treatment effects we observe. However, we strongly caution against a causal interpretation of these results. The specific subtopic that participants chose to write about is post-treatment. We randomized whether participants engaged in a conversation or short essay, and the subtopic participants focused on is likely due to treatment assignment. In other words, it is plausible that a participant thinks about one topic *because they are in the conversation condition* and they would not have thought about the subtopic had the same person been assigned to the control condition. If conversation leads people to choose subtopics that they expect will be easier to talk about with an outpartisan, then any positive effects we observe, relative to a short essay where a person may have a different decision calculus when it comes to engaging with the 2020 election topic (but recall we did not see changes in affective polarization outcomes for the control group), then that is still a causal effect of conversation, our treatment of interest.

We further addresses the idea that, if participants in the treatment (conversation) and control (short essay) spoke or wrote about the exact same topic, perhaps there would not be any treatment effect of conversation. While any converastional dynamics, including exactly what was talked about, is post-treatment, we explore how conditioning on different subtopics may affect our treatment effect estimates. For each subtopic in Figure A.6, we estimate treatment effects amongst he sample of short essays and conversations that discussed it. Table A.17 shows the results. The first model presents the treatment effect estimate of conversation, relative to writing a short essay, given the participants wrote or spoke about the candidates' actions since the 2020 election. We find a significant and similarly sized treatment effect estimate as reported in our main results. This pattern holds conditioning on each of the subtopics separately. We also find similar results if participant chose to speak or write about January 6th insurrection. Finally, we do not conduct a similar analysis for those who chose to speak or write about the inauguration as the sample size was too small.

While we focus on change in outparty affect as our main outcome in Table A.17, we conduct a similar analysis specifically to assess whether the subgroup who wrote or spoke about whether the 2020 election was administered fairly would have significant treatment effects of conversation on their perceptions of election integrity. The idea is that perhaps discussing election integrity is a necessary condition of changing those attitudes. We find that among this subgroup (N=414), the treatment effect estimate of conversation is 0.042, which is a larger effect size than in our main results. However, it is still not a significant effect (p=0.35), thus we find no evidence that conversation about election

Table A.17: Treatment Effects Conditional on 2020 Election Subtopics Discussed

	Outcome: Change in Outparty Affect				
	Candidate Actions	State of Country	Admininstration Transition	Election Fair	January 6th
Conversation	6.69*	6.83*	8.17*	6.47*	10.75*
	(2.31)	(1.71)	(3.85)	(2.29)	(4.79)
Num.Obs.	288	437	238	414	133

* p < 0.05

administration, relative to writing about that same topic, led to increases in electoral integrity.

Appendix P Willingness to Engage in Cross-partisan Conversation as a Moderator

We asked participants in the pre-treatment survey: "If you were talking with a typical member of the [Democratic/Republican] party, would you be willing to have a conversation with them about the 2020 presidential election?" The response options were definitely yes, probably yes, probably not, and definitely not. We binarized this variable into a yes or no response. In our sample, 241 (41%) said they would *not* be willing to have this kind of conversation, while 337 (59%) said they would. This is representative of our full recruitment sample where 44% indicated they would not be willing to have an interparty conversation and 56% said they would.

We assessed whether having a preference for or against interparty conversation moderated our treatment effects (a pre-registered analysis) in Table A.18. We do not find that having a preference for cross-partisan conversation moderates our main treatment effect findings.

Table A.18: Pre-Treatment Willingness to Engage in Cross-Partisan Conversation

	Outparty Affect
Conversation	6.15*
	(2.13)
Preference for Conversation	0.51
	(1.87)
Conversation × Preference for Conversation	0.17
	(2.59)
Num.Obs.	578

Note: * p < 0.05. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Appendix Q   Willingness to Engage in Cross-partisan Conversation as an Outcome

We asked respondents, both pre-treatment and post-treatment, three questions about their willingness to have future cross-partisan conversations.

First, we asked about their willingness to have a conversation with an outpartisan about the 2020 election—the same topic as the treatment. Specifically, we asked: "If you were talking with a typical member of the [Democratic/Republican] party, would you be willing to have a conversation with them about the 2020 presidential election?" The response options were a four-point Likert scale from 0 (definitely not) to 3 (definitely yes).

Second, our social polarization scale asked respondents to report their willingness to talk about "politics" with an outpartisan. The response options were on a four-point Likert scale from 0 (very unlikely) to 3 (very likely). Third, also on the social polarization scale, we asked respondents to report their willingness to talk about "sports or pop culture" with an outpartisan. The response options were on a four-point Likert scale from 0 (very unlikely) to 3 (very likely).

For all three topics, we use the change score as our outcome and estimate average treatment effects of conversation and heterogeneous treatment effects by partisanship. Table A.19 shows that conversation causes an increase in respondents' willingness to interact with an out-partisan for all three topics. Interestingly, as we abstract away from the experiment's specific topic (the 2020 election) estimated average treatment effect sizes wane, but all remain positive and significant. Consistent with our findings in the main text, we find no evidence of heterogeneous treatment effects by partisanship. We find no evidence that partisan group threat conditions the ability of conversation to decrease one's self-reported likelihood of having future cross-partisan conversations.

Table A.19: Treatment Effects of Conversation on Future Cross-Partisan Conversation

	2020 Election		Politics		Sports or Pop Culture	
Conversation	0.21*	0.23*	0.17*	0.16*	0.12*	0.19*
	(0.07)	(0.08)	(0.06)	(0.08)	(0.06)	(0.08)
Partisanship (Rep.)		0.08		0.07		0.00
		(0.06)		(0.06)		(0.05)
Conversation × Partisanship (Rep.)		-0.04		0.02		-0.13
		(0.10)		(0.10)		(0.09)
Num.Obs.	578	578	578	578	578	578

Note: * p < 0.05. The model includes blocked fixed effects to reflect the design's randomization of treatment. HC2 robust standard errors are clustered at the partnership level for individuals assigned to the conversation condition.

Appendix R   Comparison of Study to Other Cross-Partisan Conversation Research

Given space constraints in our manuscript, it was not possible to fully differentiate the novelty of our research design from extant work, so we do so in this section of the appendix. In the manuscript, we argue that our design is unique because it creates a harder test of the effectiveness of cross-partisan conversation on outparty affect than previous research. We believe this to be true because of the unique timing of our study, which was conducted immediately following Joe Biden's inauguration as President of the United States, which should have amplified partisan identity (Michelitch and Utych 2018) and feelings of threat for Republicans and reassurance for Democrats. Furthermore, this asymmetry in emotional reactions to the political context could create asymmetrical reactions to the conversations, leading one group (Democrats) to reduce their outparty animosity, while Republicans do not. In this section, we justify that these two key features of our manuscript (contentious timing against real world events and investigating heterogeneous effects by party) are indeed novel and valuable.

Contentious Timing

Previous cross-partisan conversation experiments have largely been conducted during less intense political contexts than our study and have had participants discuss topics that do not prime partisan identity, political conflict, or electoral victories (or losses) as intensely. For example, Table A.20 shows the approximate dates during which previous cross-partisan conversation, contact, or imagined contact experiments were conducted along with the topics they discussed (or imagined). Some may argue that Santoro and Broockman's (2022) paper would have a similar political climate as ours since they collected their data in 2021, as did we. While the dates of data collection are not available in the manuscript, supplementary information, or pre-registration, we can deduce that the data were collected between May 16, 2021 and August 23, 2021, based on dates of their pre-registration and revised pre-registrations. This time period was months after the November 2020 election, Joe Biden's inauguration, and the January 6th insurrection. This is not to say that partisan identity was not salient at this time. However, it was arguably less intense than the moments immediately following the formal transition of power with Joe Biden's inauguration, which is when we collected our data.

Moreover, the topic of conversation in their study did not focus on the 2020 election nor any other salient political topic (participants were asked to discuss why they were a Democrat/Republican). This is not a fault of their study – they were testing different hypotheses. But, it does highlight a way in which our design is distinct from theirs and arguably offers a significantly more contentious context in which partisans interacted.

Table A.20: Previous Cross-Partisan Contact Studies' Timing and Examination of Heterogeneous Treatment Effects by Partisanship

Study	Conversation Topic	Type of Contact	Timing	HTE by Party?
Levendusky and Stecula (2021)	Political (surprising consensus between parties) vs. non-political (best beach towns on the Jersey Shore)	In-person	November 2018; July 2019	No
Wojcieszak and Warner (2020)	Non-political (i.e. cooking show context; waiting to be seated at restaurant)	Imagined	July 2018	No
Warner and Villamil (2017)	Obama's performance as president; what America's priorities should be	Imagined	July-August 2015	No
Peña, Wolff and Wojcieszak (2021)	Immigration	Virtual Reality (no conversation)	Fall 2018	No
Huddy and Yair (2020)	Immigration	Vignette of elite contact	November 1-8 2017; July 17-25 2018	No
Rossiter (2022)	Political (gun control, immigration) vs. non-political (best meaning of life)	Written, online chat	August 2019-January 2020; September-October 2020	No
Santoro and Broockman (2022)	Thoughts on the perfect day, why they like their inparty, why they dislike outparty	Video Chat	Summer-Fall 2021 (pre-registrations posted 5/16/21 and 8/23/21)	Yes (SI Figure S7; Not pre-registered)
Combs et al. (2022)	Immigration, gun control	Written, online chat	February 2020	Yes (Fig 4; Not pre-registered)
Kalla and Broockman (2022)	Immigration (Studies 1 and 2); Supporting Biden over Trump (Study 3)	Door to door canvassing (Study 1); phone (Studies 2 and 3)	Study 1: Winter 2019-2020; Study 2: summer 2020; Study 3: during 2020 election	No
Fishkin et al. (2021)	Immigration, economy, health care, environment, foreign policy	In-person, moderated deliberative poll ("America in One Room")	September 19-22, 2019	Results presented by party, but not hypothesized
Baron et al. (2021)	Stereotype, fishbowl questions, and how can we contribute exercises	Moderated in-person workshop	Spring 2020	Pre-registered HTE by party, but did not specify direction: "HTE1:The magnitude of the treatment effect on our primary outcome, affective polarization, will vary with party ID" (Pre-analysis plan p. 12)

Heterogeneous Treatment Effects by Party

Table A.20 above also highlights whether the manuscripts reported treatment effects by party in the main text. One article and two unpublished manuscripts do so. The first is Santoro and Broockman (2022), which makes one brief reference to a figure in the supplementary information: "(In response to feedback, we conducted a post hoc test, which found similar effects on effects toward outparty voters for Democratic and Republican respondents; see fig. S7.)" (p 6). This analysis was not included in the author's pre-registration and was not part of the authors' original theory.[3]

The second manuscript to do so is Combs et al. (2022). The authors report conditional treatment effects by party in Figure 4 of the main text. Unlike our findings, this study find that Republicans have significant positive treatment effects of cross-partisan conversation (p<.01), but the treatment effects on Democrats are small and cannot be distinguished from zero (p=.48). Despite these different point estimates, the authors are not powered to find Democrats and Republicans responded differently to cross-partisan converastion (p=.14). Importantly, while this study was pre-registered, heterogeneous treatment effects by partisanship was not included as a pre-registered analysis nor was it a part of the authors' theory.[4]

Moreover, an unpublished manuscript (Baron et al. 2021) pre-registers a test for variation based on party, but does not specify a direction, writing: "We will test whether the DW and SBD workshops are more effective among Democrats or Republicans; and whether the mixed party and inparty-only workshops are more effective among Democrats or Republicans. But we do not have strong priors about these heterogeneous treatment effects, and so do not specify their direction" (pg 12).[5]

In contrast to the three studies discussed above, heterogeneous treatment effects between Democrats and Republicans is a core component of our inquiry, was part of our theoretical development, pre-registration, and tests presented in the manuscript. While we are not the first to analyze heterogeneous treatment effects of cross-partisans conversation for Republicans and Democrats, we believe our paper to be the first to directly theorize about and test these effects as a core component of the manuscript.

---

[3]Pre-registrations for the Santoro and Broockman (2022) studies: https://aspredicted.org/blind.php?x=43v3vh, https://aspredicted.org/blind.php?x=ud7a62.

[4]Pre-registrations for the Combs et al. (2022) study: https://osf.io/g97z5/.

[5]Pre-registration for the Baron et al. (2021) study: https://osf.io/w8kpe

References

Barlow, Fiona Kate, Stefania Paolini, Anne Pedersen, Matthew J Hornsey, Helena RM Radke, Jake Harwood, Mark Rubin and Chris G Sibley. 2012. "The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice." *Personality and Social Psychology Bulletin* 38(12):1629–1643.

Baron, Hannah, Robert Blair, Donghyun Danny Choi, Laura Gamboa, Jessica Gottlieb, Amanda Lea Robinson, Steven Rosenzweig, Megan Turnbull and Emily A West. 2021. "Can Americans Depolarize? Assessing the effects of reciprocal group reflection on partisan polarization.".

Combs, Aidan, Graham Tierney, Brian Guay, Friedolin Merhout, Christopher A Bail, D Sunshine Hillygus and Alexander Volfovsky. 2022. "Anonymous Cross-Party Conversations Can Decrease Political Polarization: A Field Experiment on a Mobile Chat Platform.".

Connors, Elizabeth C and Christopher Howell. 2022. ""You Need to Calm Down": How Tone Shapes Political Discussions." Working Paper.
URL: *http://www.elizabethchaseconnors.com/images/Tone.pdf*

Druckman, James N, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John Barry Ryan. 2022. "(Mis) estimating affective polarization." *The Journal of Politics* 84(2):000–000.

Fishkin, James, Alice Siu, Larry Diamond and Norman Bradburn. 2021. "Is deliberation an antidote to extreme partisan polarization? Reflections on "America in one room"." *American Political Science Review* 115(4):1464–1481.

Huddy, Leonie. 2013. "From group identity to political cohesion and commitment.".

Huddy, Leonie and Alexa Bankert. 2017. Political partisanship as a social identity. In *Oxford research encyclopedia of politics*.

Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." *American Political Science Review* 109(1):1–17.

Huddy, Leonie and Omer Yair. 2020. "Reducing Affective Polarization: Warm Group Relations or Policy Compromise?" *Political Psychology* .

Kalla, Joshua L and David E Broockman. 2022. "Voter Outreach Campaigns Can Reduce Affective Polarization among Implementing Political Activists: Evidence from Inside Three Campaigns." *American Political Science Review* pp. 1–7.

Levendusky, Matthew S and Dominik A Stecula. 2021. *We Need to Talk: How Cross-Party Dialogue Reduces Affective Polarization.* Cambridge University Press.

Long, Jacob A, William P Eveland Jr and Michael D Slater. 2019. "Partisan media selectivity and partisan identity threat: The role of social and geographic context." *Mass Communication and Society* 22(2):145–170.

Mason, Lilliana. 2016. "A cross-cutting calm: How social sorting drives affective polarization."

Public Opinion Quarterly 80(S1):351–377.

Peña, Jorge, Grace Wolff and Magdalena Wojcieszak. 2021. "Virtual Reality and Political Outgroup Contact: Can Avatar Customization and Common Ingroup Identity Reduce Social Distance?" *Social Media+ Society* 7(1):2056305121993765.

Pérez, Efrén O. 2015. "Xenophobic rhetoric and its political effects on immigrants and their co-ethnics." *American Journal of Political Science* 59(3):549–564.

Pew Research Center. 2020. "Sharp Divisions on Vote Counts, as Biden Gets High Marks for His Post-Election Conduct.".

Rossiter, Erin. 2022. "The Consequences of Interparty Conversation on Outparty Affect and Stereotypes." Working paper. https://erossiter.com/files/conversations.pdf.

Santoro, Erik and David E Broockman. 2022. "The promise and pitfalls of cross-partisan conversations for reducing affective polarization: Evidence from randomized experiments." *Science advances* 8(25):eabn5515.

Warner, Benjamin R and Astrid Villamil. 2017. "A test of imagined contact as a means to improve cross-partisan feelings and reduce attribution of malevolence and acceptance of political violence." *Communication Monographs* 84(4):447–465.

Wojcieszak, Magdalena and Benjamin R Warner. 2020. "Can interparty contact reduce affective polarization? A systematic test of different forms of intergroup contact." *Political Communication* pp. 1–23.

Young, Lori and Stuart Soroka. 2012. "Affective news: The automated coding of sentiment in political texts." *Political Communication* 29(2):205–231.