Online Appendix

Contents

A Additional Tables and Figures **2**

B Ethical Considerations **4**

C Additional Details on Survey and Campaign Context **5**

C.1 Sample Demographics 8

D Regression and Paired t-test Numerical Results **10**

D.1 Heterogeneous Treatment Effects by Prior Experience with Perspective-Getting Canvassing 13

D.2 Heterogeneous Treatment Effects by Number of Conversations 13

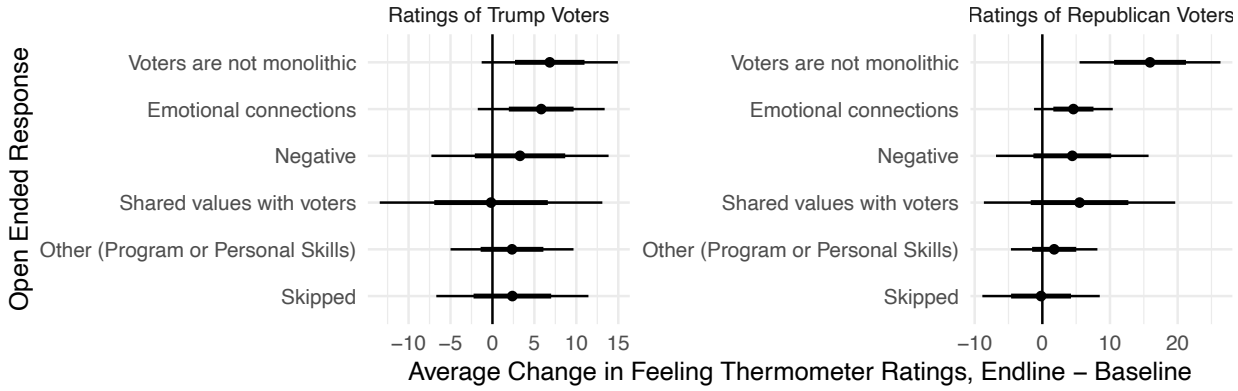
E Endline Survey Attrition **14**

F Qualitative Evidence from Study 3 **16**

G References for Online Appendix **17**

A Additional Tables and Figures

Figure A1: Study 3: Changes in Affective Polarization, by Coding of Open-Ended Response



Notes: See Table A14 for examples of each category.

Table A1: Prior research on inter-partisan contact and reductions in affective polarization

(a) Studies of in-person interpersonal conversations conducted by activists, or as part of campaigns

No prior literature; only this paper

(b) Studies of in-person interpersonal conversations between partisans in field settings

Citation	Intervention	Sample
Levendusky and Stecula (2021)	Read article about common ground between parties; discuss article in heterogeneous group that includes outpartisans	Metropolitan Philadelphia lab-in-the-field experiment
Baron et al. (2021)	Braver Angels workshop on how to bridge political divides (field experiment)	College students

(c) Online interpersonal conversations between partisans

Citation	Intervention	Sample
Rossiter (2021)	Online text-based discussions about issues	Democrats and Republicans on Mechanical Turk

(d) Non-exhaustive examples of other studies not involving interpersonal conversation or contact

Citation	Intervention	Sample
Wojcieszak and Warner (2020)	Read a mock news story (Study 2); Imagined contact (Study 3)	Nationally representative online survey experiment with pure independents excluded.
Huddy and Yair (2021)	Read a mock news story	Democrats and Republicans on Mechanical Turk
Levendusky (2018*a*)	Prime American identity with a news article	Nationally representative online survey experiment
Levendusky (2018*b*)	Inducing partisan ambivalence; self-affirmation	Nationally representative online survey experiment.

Notes: There are a number of studies that test approaches for reducing affective polarization and Table A1d does not attempt to review them exhaustively; it merely provides examples. There are also other studies that do not examine impacts on affective polarization, such as Klar (2014), that we do not include here.

B Ethical Considerations

The authors declare that the human subjects research in this article was reviewed by the Yale University Human Subjects Committee and was determined to be exempt. The authors affirm that this article adheres to the APSA's Principles and Guidance on Human Subjects Research. Participants provided informed consent.

C Additional Details on Survey and Campaign Context

The three programs where we embedded these surveys are summarized in Table A2.

Table A2: Overview of Studies

Study	Topic of Canvassing Program	Baseline Survey Dates	Endline Survey Dates	Median Days Between Last Conversation and Endline Survey
Study 1	Immigration	Oct. 31, 2019 - Feb. 8, 2020	Jan. 12 - March 7, 2020	23
Study 2	Immigration	June 18 - July 16, 2020	August 2 - 27, 2020	7
Study 3	2020 Presidential Election	Oct. 2 - 20, 2020	Oct. 28 - Nov. 3, 2020	Unknown

Study 1 was administered to the canvassers who implemented the program studied in Kalla and Broockman (Forthcoming, Study 2). Canvassers had door-to-door canvass conversations in California and Tennessee in an attempt to reduce exclusionary attitudes towards undocumented immigrants. 48% of the conversations were with voters who identified as Republican or conservative. 26 political activists completed the baseline survey between October 31, 2019 - February 8, 2020. Canvassing then took place from January 12 - March 7, 2020. 23 political activists completed the endline survey between March 22-29, 2020 (88% reinterview rate).[9] A median of 23 days elapsed between political activists' last conversation and when they completed the endline survey. On average, canvassers had 6 perspective-getting conversations in this study.

After taking the baseline survey, canvassers in Study 1 were trained in how to have these conversations. Canvassers received both an initial training as well as ongoing training and feedback. The trainings focused on providing canvassers with the skills to listen to and ask questions of voters in a non-judgmental manner that would elicit narratives from voters about

[9] A pre-analysis plan is available at https://osf.io/c5e94/?view_only=e38a43b662cc4a4ba809398d76853be5. Other than the feeling thermometers, we do not report the other outcome measures. No significant movement was detected on them.

their experiences. Trainings often involved role play, viewing video of past canvass conversations, and following a more experienced canvasser for a day.

Study 2 surveyed the canvassers who implemented the program studied in Kalla and Broockman (Forthcoming, Study 3). Canvassers had phone conversations in California and Tennessee in an attempt to reduce exclusionary attitudes towards undocumented immigrants.[10] 44% of the conversations were with voters who identified as Republican or conservative. 28 political activists completed the baseline survey between June 18 - July 16, 2020. Calling then took place between July 6 - August 3, 2020. 23 political activists completed the endline survey between August 2-27, 2020 (82% reinterview rate).[11] A median of 7 days elapsed between political activists' last phone call and when they completed the endline survey. On average, canvassers had 9 perspective-getting conversations in this study. Study 2 involved a similar training as Study 1.

Study 3 took place in the context of a persuasion program conducted by a political organization, People's Action. People's Action is a multi-state grassroots political organization that made phone calls during the 2020 presidential campaign to persuade voters in targeted swing states to support Biden over Trump for president. As part of their efforts, political activists from People's Action had phone conversations with voters the organization estimated to be Republican-leaning in Arizona, Colorado, Michigan, Minnesota, North Carolina, New Hampshire, Pennsylvania, and Wisconsin. 161 political activists completed the baseline survey between October 2-20, 2020 and 104 political activists completed the endline survey between October 28-November 3, 2020 (65% reinterview rate).[12] On average, canvassers had 23 perspective-getting conversations in this study.

The training program for Study 3 focused on building core skills of nonjudgmental listening,

[10] This also includes an additional program in Fresno California that did not have the difference-in-differences design reported in Kalla and Broockman (Forthcoming).

[11] This study was not pre-registered. Instead, we follow the same analysis plan as Study 1. The surveys were identical across both studies.

[12] A pre-analysis plan is available at https://osf.io/7em2a/?view_only=c40755b3b86b43f58aca65df5397d015. We do not have data on when political activists called, so we are unable to calculate the average number of days between the last call and when political activists took the survey.

sharing stories with emotional resonance, compassionate curiosity, and holding space for voters to process conflict. The training and orientation for canvassers focused on staying grounded in sharing lived experiences and emphasized the importance of not arguing with voters. There was also an attempt to integrate trauma-informed practices to allow canvassers to process difficult conversations. The training also included an overview of the script and example conversations.

C.1 Sample Demographics

Table A3: Study 1 Summary Statistics

	mean	sd
Age	24.26	4.86
Female	0.74	0.45
Latino	0.30	0.47
Immigrant	0.22	0.42
Very Liberal	0.48	0.51
Liberal	0.22	0.42
Moderate	0.30	0.47
Conservative	0.00	0.00
Very Conservative	0.00	0.00
Prior Campaign Experience	0.57	0.51
Prior Perspective-Getting Experience	0.30	0.47
Observations	23	

Table A4: Study 2 Summary Statistics

	mean	sd
Age	20.96	3.44
Female	0.70	0.47
Latino	0.39	0.50
Immigrant	0.48	0.51
Very Liberal	0.30	0.47
Liberal	0.43	0.51
Moderate	0.26	0.45
Conservative	0.00	0.00
Very Conservative	0.00	0.00
Prior Campaign Experience	0.57	0.51
Prior Perspective-Getting Experience	0.35	0.49
Observations	23	

Table A5: Study 3 Summary Statistics

	mean	sd
Age	31.60	13.55
Male	0.44	0.50
Female	0.48	0.50
White	0.75	0.44
African American	0.15	0.36
College Grad	0.68	0.47
Democrat	0.60	0.49
Independent or Other	0.40	0.49
Republican	0.00	0.00
Very Liberal	0.74	0.44
Liberal	0.21	0.41
Moderate	0.05	0.21
Prior Campaign Experience	0.65	0.48
Prior Perspective-Getting Experience	0.17	0.38
Observations	104	

D Regression and Paired t-test Numerical Results

Following our pre-analysis plans, the following tables create a stacked dataset with each row corresponding to a separate survey response from a political activist. We then regress each outcome on an indicator for survey wave and use political activist fixed effects to estimate the within-political activist change from the baseline to the endline. As can be seen in Table A9, this produces identical numerical results to the paired t-tests we report in the main text for simplicity.

Table A6: Study 1 Regression Results

	Trump Supporters	Trump Opponents	R Voters	D Voters	R Pols	D Pols	Dis. Imm.	Agree Imm.
Change	9.565***	1.870	3.696	0.174	3.087	-0.174	4.957	0.130
	(3.163)	(4.239)	(3.096)	(3.291)	(3.244)	(3.644)	(4.772)	(2.685)
Const.	16.48***	74.13***	35.04***	71.17***	26.30***	67***	29.43***	85.78***
	(2.236)	(2.997)	(2.189)	(2.327)	(2.294)	(2.576)	(3.374)	(1.899)
N	46	46	46	46	46	46	46	46

Standard errors in parentheses

All models include canvasser fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Study 2 Regression Results

	Trump Supp.	Trump Opp.	R Voters	D Voters	R Pols	D Pols	Dis. Imm.	Agree Imm.
Change	2.957	-5.174	2.130	-4.348	1.174	-0.130	2.826	-4.783
	(2.753)	(3.271)	(3.337)	(3.056)	(2.980)	(3.591)	(5.591)	(5.670)
Const.	14.87***	85.52***	30.13***	76.65***	20.35***	67.22***	28.74***	85.96***
	(1.947)	(2.313)	(2.360)	(2.161)	(2.107)	(2.539)	(3.954)	(4.010)
N	46	46	46	46	46	46	46	46

Standard errors in parentheses

All models include canvasser fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Study 3 Regression Results

	R Voters	Trump Voters	R Pols.	D Voters	Not Trump Voters	D Pols.
Change	4.221**	3.683*	1.952	0.712	-1.442	0.394
	(1.759)	(1.896)	(1.330)	(1.378)	(2.113)	(1.490)
Const.	29.65***	18.83***	10.10***	76.67***	83.84***	60.67***
	(1.244)	(1.341)	(0.941)	(0.975)	(1.494)	(1.054)
N	208	208	208	208	208	208

Standard errors in parentheses

All models include canvasser fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Paired t-tests

Outcome	**Study**	**Difference in Means**	**Std. Error**	p**-value**
People who support Donald Trump	1	9.57	3.16	0.01
	2	2.96	2.75	0.29
Trump voters	3	3.68	1.90	0.05
Trump supporters/voters	Pooled	4.47	1.47	$<$0.01
Republican voters	1	3.70	3.10	0.25
	2	2.13	3.34	0.53
	3	4.22	1.76	0.02
	Pooled	3.82	1.40	$<$0.01

D.1 Heterogeneous Treatment Effects by Prior Experience with Perspective-Getting Canvassing

Table A10: Heterogeneous Treatment Effects by Prior Experience with Perspective-Getting Canvassing

	Trump Supporters	Trump Supporters	R Voters	R Voters
Change	4.359**	4.879	4.590***	1.091
	(1.681)	(3.039)	(1.569)	(3.071)
Const.	17.923***	17.636***	30.299***	31.454***
	(1.189)	(2.149)	(1.110)	(2.172)
Prior Experience?	No	Yes	No	Yes
N	234	66	234	66

Standard errors in parentheses

All models include canvasser fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D.2 Heterogeneous Treatment Effects by Number of Conversations

Table A11: Heterogeneous Treatment Effects by Number of Conversations (Studies 1 and 2)

	Trump Supporters	Trump Supporters	R Voters	R Voters
Change	3.850	10.824**	0.150	5.294
	(3.020)	(3.767)	(1.980)	(4.290)
Const.	17.050***	13.882***	32.800***	35.529***
	(2.136)	(2.664)	(1.400)	(3.033)
Above Median #?	No	Yes	No	Yes
N	40	34	40	34

Standard errors in parentheses

All models include canvasser fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Endline Survey Attrition

One potential threat to validity could be that respondents to the endline survey are distinct from the broader set of canvassers in these three studies who took the baseline survey. To assess this, we regress an indicator for whether or not a canvasser participated in the endline survey on a series of baseline measures. The results are shown in Table A12. This was not a pre-registered analysis.

Table A12: Predictors of responding to the endline survey

	Responded to Endline Survey
Female	-0.020 (0.068)
Person of Color	-0.185 (0.076)
Age	0.008 (0.003)
Very Liberal	0.049 (0.112)
Liberal	0.143 (0.116)
Feeling therm: Dem voters	0.001 (0.002)
Feeling therm: Trump voters	-0.001 (0.002)
Prior campaign experience	-0.041 (0.068)
Constant	0.761 (0.191)
Study Fixed Effects	Yes
N	215

Note: Robust HC3 standard errors in parentheses.

There are two statistically significant predictors: being a person of color and age. However, substantively, the average age of canvassers who took the endline survey (26.29), is only modestly different from canvassers who did not (28.84). Similarly, for being a person of color (40% among those who responded to the baseline survey vs. 48%). Overall, there do not appear to be meaningful differences between canvassers who did and did not take the endline survey, especially when examining baseline political attitudes.

Based on a suggestion from an anonymous reviewer, for Studies 1 and 2, we can also test whether patterns of attrition are different depending on the types of voters the canvassers spoke with. Each conversation began with the canvasser asking the voter how they felt on the political issue being discussed, ranging from 0 (most conservative response) to 10 (most progressive

response). For each canvasser, we calculate the average of these ratings collected at the start of the conversation.

To test whether attrition is predicted by the average ratings of the conversations a canvasser had, we regress an indicator for responding to the endline survey on the average rating and include a study fixed effect. In this regression, we exclude the 4 Study 1 subjects and 13 Study 2 subjects for whom we are missing this data. Because these studies were not originally conducted with this analysis in mind, this data was not always systematically collected. For example, in Study 2, this data was not collected at one of the study sites.

The regression results are shown in Table A13. Overall, we find that the average start rating of conversations is not a significant predictor of attrition.

Table A13: Predictors of responding to the endline survey by conversation quality (Studies 1 & 2)

	Responded to Endline Survey
Average Rating	0.014 (0.016)
Constant	0.845 (0.124)
Study Fixed Effects	Yes
N	36

Note: Robust HC3 standard errors in parentheses.

F Qualitative Evidence from Study 3

Table A14: Open-ended coding from Study 3, response to "What aspects of this project most surprised you?"

Category	Frequency	Example
Humanizing emotional connections	26%	"How willing people are to talk to a stranger on the phone. The thing I'll take away from this job most of all is that everyone wants someone to ask them how they're doing, and to really listen, and pay attention. A lot of people hang up on you when you ask them about their life, it's true, but a shocking number of people also desperately want to build a connection with you as long as you're brave enough to ask them how they are, and patient enough to REALLY listen to their answer. Everyone is always just reaching out and looking for human connections"
Individuated, non-monolithic	12%	"I was expecting more similarities among the ideals held by undecided voters, but it really is true that no two voters are the same and it really challenged my ideas of what a trump-leaning voter would sound like."
Shared values	6%	"The humanity of Trump voters most surprised me in this project. Even while being able to identify how both political parties demonize the other, I still fell victim to viewing Trump supporters in this way. I realized, however, when encountering them on the phones, that most Trump supporters are not that different from ourselves—they want the same things and hurt for the same reason. Most are not hateful in the way that they are portrayed to be, and it was refreshing to make emotional connections with some of them."
Other (Program)	28%	"My coworkers. I didn't expect the environment to be quite this good."
Other (Personal skills)	7%	"The emotional labor that can go into this work. Having Leaders like [NAMES] that were transparent about the work and their willingness to supported us in any way as we did this work. 10 years in the work force and this was the first job I felt truly valued for my individuality skill set and still pushed to developed those skills rather than pushed to conform to a work style that was uncomfortable to me."
Negative	7%	"The surprising vitriol and racism from certain individuals"
Skipped	15%	n/a

Notes: $n = 104$. *Each response coded into only one category. The authors individually coded each statement, blind to each other's coding. The authors agreed on 90% of the codings after this initial pass. For the remainder, the authors discussed an appropriate coding.*

G References for Online Appendix

References for Appendices

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." Available at https://osf.io/3x7z8.

Huddy, Leonie and Omer Yair. 2021. "Reducing Affective Polarization: Warm Group Relations or Policy Compromise?" *Political Psychology* 42(2):291–309.

Klar, Samara. 2014. "Partisanship in a social setting." *American Journal of Political Science* 58(3):687–704.

Levendusky, Matthew S. 2018*a*. "Americans, not partisans: Can priming American national identity reduce affective polarization?" *Journal of Politics* 80(1):59–70.

Levendusky, Matthew S. 2018*b*. "When efforts to depolarize the electorate fail." *Public Opinion Quarterly* 82(3):583–592.

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

Rossiter, Erin. 2021. "The Consequences of Interparty Conversation on Outparty Affect and Stereotypes." Available at https://erossiter.com/files/conversations.pdf.

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.