

Evaluating an In-Person Depolarization Program and Testing a Light-Touch Enhancement

Mark J. Brandt

Department of Psychology, Michigan State University

Jerad Morey

Respectful Conversations, Minnesota Council of Churches

Correspondence can be addressed to Mark Brandt (m.j.brandt@tilburguniversity.edu) and Jerad Morey (j.morey@mnchurches.org)

Acknowledgments: Funding for this project was provided by Stanford University's Strengthening Democracy Challenge. We thank Stanford's Polarization and Social Change Lab for feedback on a prior version of this project. We thank all of the facilitators and participants of the MCC Respectful Conversation events. The first author is now affiliated with Department of Social Psychology, Tilburg University (Netherlands).

Conflict of Interest Statement: Jerad Morey is Director of Civic Dialogue at Minnesota Council of Churches, in which role he stewards the Respectful Conversations project.

Transparency Statement: The experiment was preregistered and includes three updates with different inclusion criteria as more events became available and other events were rescheduled. The preregistration, including all updates, is available here:

https://osf.io/fr4y3/?view_only=595eb87d613f47ffb2574b8166f23661. Data from the experiment and code to reproduce the figures and analyses in this manuscript are available here: https://osf.io/hke9w/?view_only=71362e6c532d402d86c200d80461943f.

Author contributions

Mark Brandt: Conceptualization, data curation, formal analysis, methodology, software, visualization, writing – original draft

Jerad Morey: Conceptualization, data curation, funding acquisition, investigation, methodology, resources, writing – review and editing

In press at Social Psychology

Abstract

We evaluated whether an in-person, multi-hour conversation intervention reduced animosity toward people with different (political) attitudes. We also experimentally tested whether stacking a brief video intervention highlighting sympathetic narratives onto the intensive intervention boosted its ability to reduce animosity. Across events ($N = 209$, Events $k = 10$), participants reported less animosity and more empathy following the events compared to before, suggesting the intervention is effective. Different than preregistered, this effectiveness was not significantly affected by the light-touch video treatment. The program improved intergroup relations, but we did not boost the intensive treatment's effects with the addition of a light-touch video treatment.

(100 words)

Evaluating an In-Person Depolarization Program and Testing a Light-Touch Enhancement

People are less likely to like people they disagree with (Byrne, 1969), an effect so consistent some have called it a psychological law (Byrne & Griffitt, 1968). When a person or group is perceived to hold different political views, people express animosity towards that group and are more likely to discriminate against that group (e.g., Brandt & Crawford, 2020; Crawford et al., 2017; Gift & Gift, 2015; Whitt et al., 2021; Woitzel & Koch, 2023). In this way, disagreements can degrade the social fabric of our communities. Community members and scholars are both interested in how to reduce the social fraying caused by disagreement (Hartman et al., 2022). It could have both practical benefits (more social cohesion) and theoretical insight (how to alter a psychological “law”). Here we test an intensive in-person, conversation-based intervention designed to help people manage and obtain perspective about people they disagree with in their community. As part of our evaluation, we also test if we can improve upon this program by stacking a light touch video intervention on it in a preregistered field experiment. We do this within the context of the Respectful Conversations program run by the Minnesota Council of Churches (hereafter MCC Respectful Conversations).

Approaches to Interventions

Light-Touch Experimental Literature

Scholars have identified factors that might be used to reduce the negative interpersonal consequences of disagreement (Hartman et al., 2022; Voelkel et al., 2024). For example, imagining being included by people on the other side (Voelkel et al., 2021), receiving misperception corrections (Ahler & Sood, 2018; Mernyk et al., 2022; Moore-Berg et al., 2020), and reading about warm interactions between opposing political elites (Huddy & Yair, 2021) have all been found to reduce political animosity in online and survey experiments. One integration of the literature suggests that those interventions that reduce misperceptions, encourage empathy, or a combination of both (as in cases of ideal contact) are the most successful (Hartman et al., 2022).

The typical approach to testing and developing interventions in the academic literature is to use tightly controlled lab or survey experiments with diverse, if not representative samples of participants (for a study including many such interventions, see Voelkel et al., 2024). These choices maximize internal validity but simultaneously constrain the types of studies that are conducted. For example, conducting a survey experiment on a representative sample requires that the intervention is short and easily presented in the survey context (often in an online questionnaire). Such designs can identify causal effects

to inform psychological theory, but it is not clear that these designs' effects are robust (e.g., Dias et al., 2024; Holliday et al., 2025; Voelkel et al., 2024). Consistent with this, a meta-analysis of the prejudice reduction literature suggests that most of the field's methods produce weak effects that are unlikely to last (Paluck et al., 2021). Although this meta-analysis did not include studies about attitude disagreement, we suspect that this lesson is generalizable (e.g., Dias et al., 2024; Holliday et al., 2025; Voelkel et al., 2024).

Intensive Interventions

One response to the lack of durable effects of light-touch interventions is to construct more intensive interventions. A growing number of civic-minded organizations interested in reducing polarization and bridging political (and related) differences have developed programs to do just this. The precise goals and methods of the bridging community differ depending on the specific organization, but in many cases these practitioners are developing programs aimed at helping people connect across differences. Just as with academic approaches, these programs often aim to reduce misperception and foster empathy. In some cases, they also aim to teach skills for discussing high conflict issues. For example, the organization Braver Angels developed Red/Blue workshops. These workshops aim to integrate information about the policies and perspectives of political outgroups with empathy building in a full-day event including structured and moderated exercises and discussions. Essential Partners, formerly the Public Conversations Project, developed reflective structured dialogue in the early 1990s. It is a model of empathy-building across divides in deeply-held values using participant engagement based on family therapy practices.

There are relatively few evaluations of these programs. However, there are some. For example, an experimental evaluation of Braver Angels Red/Blue in-person workshops found that the program reduced animosity among college students (Baron et al., 2025). An evaluation of the 10-part Narrative 4 depolarization program among high school students also finds that the program impacts affective polarization (Kubin et al., 2026). The results from these two programs are consistent with other findings from high impact studies. For example, when activists have in-depth conversations with political opponents as they canvas door-to-door, their political animosity decreases compared to prior to the conversations (Kalla & Broockman, 2022). Similarly, it is consistent with a study of a large, multi-day deliberative polling event. In this event, people discussed five polarizing topics and policies with people from a range of backgrounds to understand people's opinions when they have the time and information to fully consider an issue. A by-product of these discussions was reduced political animosity (Fishkin et al., 2021).

Not all evidence suggests that in-depth community interventions will have an impact. In two experiments, some scholars found that discussions with the political outparty only reduced affective polarization when the topics were *not* about politics (Santoro & Broockman, 2022). There are also inconsistencies with programs run by civic organizations. Braver Angels also administer two shorter interventions with similar depolarization goals for online video conferencing (e.g., with Zoom). These short and more accessible interventions expand the type of people who can participate. The test of these interventions revealed mixed results depending on the analytic focus. One analysis (Blair et al., 2025b) compared the effectiveness of the two interventions separately. Participants completed a workshop called “Skills for Bridging the Divide” where they learn skills for constructive engagement with their outgroup. These participants were compared to participants who completed a workshop called “Depolarize Within” where they learn skills to recognize their own animosity, talk about the outgroup with non-polarizing language, and challenge the ingroup when they use polarizing language. The scholars found that the “Depolarize Within” workshop was the only workshop that reduced political animosity compared to the control condition. However, another analysis using the same data focused on the effectiveness of the interventions for subsamples from different recruitment sources (Blair et al., 2025a). In this second analysis, the authors pooled the two programs and split the total sample by the recruitment source. Some participants were recruited from a national panel and another group of participants were recruited from the Braver Angels listserv. The scholars reasoned that the people on the listserv are already interested in depolarization and so may respond differently than people from the national sample. The scholars then pooled the two programs for analysis and found that participating in a workshop reduced out-party animosity compared to the control condition, but only for people from the national sample. This nuanced pattern of results across these subsample analyses suggest that in-depth, community-based interventions have potential, but may not always be successful. They require further testing.

The Program We Evaluate: MCC Respectful Conversations

We contribute to the evaluation of practitioner-lead programs by testing and building upon the MCC Respectful Conversations program. MCC Respectful Conversations are facilitated, structured dialogue events designed to produce empathy across divides, strengthen community relationships, decrease polarization and increase peacebuilding capacity among participants by facilitating interactions between people who disagree with each other (Morey, 2013). Notably, the dimensions of disagreement may or may not follow a stereotypical liberal/conservative dimension. The program does not specifically ask communities to frame their disagreements in these terms and instead works with the communities to discern the best way to frame such disagreements (e.g., “The Election”,

“Bridging Cultural Differences”). Following the discernment, volunteers from the host community are trained by a Lead Facilitator to be Table Facilitators. They learn the structure of the dialogue, they experience it as participants, and they gain skills in reframing statements and gotcha-style questions into questions of genuine interest focused on narrative-sharing.

At a MCC Respectful Conversation event, the Lead Facilitator and now-trained Table Facilitators lead a large community group through a dialogue process. The process begins with a meal. Content about potential conflict and behavioral guidelines is delivered to the whole group, as well as to small groups at tables. After this, within the small groups at tables, people take turns going around the table and share their perspective on the topic, as well as asking and responding to questions. Then there is a sort of conversational free-for-all about the topic, followed by another round of reflection on the experience. The event ends with a closing response round as a large group and evaluations.

Across the event, the structure and content are designed to increase awareness of and slow down the experience of the body’s physiological response to threat when it is confronted by ideas that challenge worldview (e.g., van der Lee et al., 2024). By using tools like silence, structured questions, and joyful animation participants are more able to be fully present with each other (see also Gower et al., 2020). Questions posed are designed to evoke personal narrative, values sharing, and emotional revelation in way that moves the discussion away from a binary right/wrong topic area into one where participants experience more curiosity about each other. This is consistent both with scholarship finding that considering multiple perspective can help reduce prejudices (e.g., Kalla & Broockman, 2023; Shih et al., 2009), as well as scholarship finding that reducing misperceptions reduces political animosity (e.g., Lees & Cikara, 2021; Moore-Berg et al., 2020). In combination, MCC Respectful Conversations aims to fortify communities through enhanced relationships that result from higher-quality conversations about disagreements.

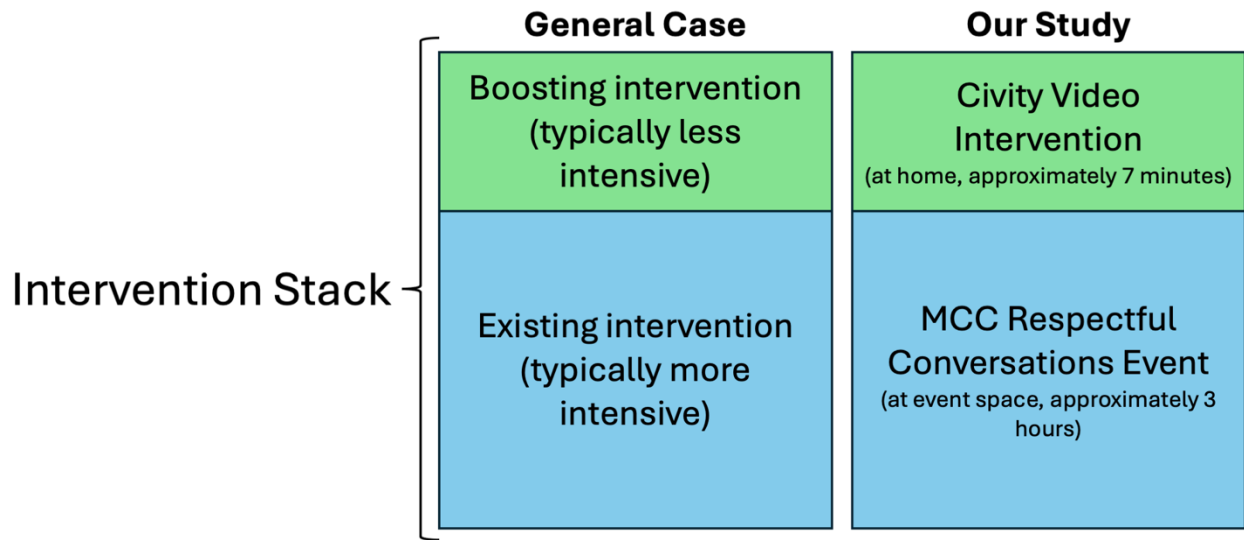
Intervention Stacking

Both scholars and practitioners have developed methods that potentially reduce animosity. A challenge is how to best integrate the findings from scholars that test less intensive interventions, with the more intensive programs developed by practitioners. One strategy already in use is to take the theoretical principle from scholarly work and embed it into the program. For example, the Red/Blue workshops noted above and MCC Respectful Conversations both take social psychological work on empathy (Batson & Ahmad, 2009) and embed it into their programs.

In this project, we engaged in scholar-practitioner collaboration, bringing together expertise in the research literature with expertise from working in the field. We tested an additional strategy that we call intervention stacking (see Figure 1).¹ This intervention stacking strategy pairs an existing and typically more intensive intervention with a boosting intervention that is typically less intensive, with roots in the research literature. This is different from the intervention approach described in the prior paragraph. That approach embeds a theoretical principle from the research literature into a larger intervention. In our approach, we are taking a light touch intervention that would usually be deployed on its own and stack it on a more intensive intervention with the aim of boosting the more intensive intervention. That means that the boosting intervention is not fully integrated into the existing intervention and is instead deployed in a very similar way as it was used in the scholarly literature. *If this strategy is successful*, it would provide an easy method for practitioners to improve on their programs using methods from the scholarly literature. In other words, by stacking the interventions, the less intensive intervention may be an easy method to improve the outcomes of the more intensive intervention.

Although this approach is born from a practical concern of simply improving interventions, there are reasons to believe that it may work. In particular, assuming that the boosting intervention works as intended (e.g., increasing empathy, reducing misperceptions, increasing openness to others) in the short-term, this primes people to be accepting and less defensive of the contents of the more intensive intervention. As an analogy, in some kinds of restaurants and cuisines, there is a small bite or taste (cf. palate cleanser, transitional course, sorbet) that is used to prepare the tastebuds for the main course that follows. Our idea is that a boosting intervention can serve as the sorbet for the intensive intervention main course.

¹ Scholar-practitioner collaborations are collaborations between scholars in academia and practitioners working in field settings. They are rare in this area (see Hartman et al., 2022). For other scholar-practitioner collaborations see Blair et al., 2025a, 2025b; Baron et al., 2025.

Figure 1. *Intervention stack in the general case and in our study***Boosting Intervention**

We used the MCC Respectful conversations as our existing and more intensive intervention. To identify a boosting intervention, we reviewed the interventions tested as part of a depolarization mega-study that tested 24 different interventions in an online sample (Voelkel et al., 2024). This project solicited potential interventions that could be used in a survey experiment to decrease political animosity and increase support for democratic norms. After identifying promising interventions, they experimentally tested 25 interventions in a large sample of US Americans. They found that 23 interventions reduce political animosity, 6 reduce support for undemocratic practices, and 5 reduce support for political violence. Of those successful interventions, our team discussed which interventions both fit well with the goals of MCC Respectful Conversations and would make the most sense to incorporate into the typical MCC Respectful Conversations process.

Based on our discussions, we chose a short, seven-minute video intervention that stresses the importance of individual stories (called “sympathetic personal narratives” in Voelkel et al., 2024) developed by the organization Civity (Kopell, Strand, Baleria, & Fiorella, 2024). In the videos, individuals talk about themselves and their experiences, often complicating viewers’ ideas about the people based on their surface-level characteristics. Following these, participants then watch a video that discusses the value of connecting with people who are different and seeing them as members of their community. The focus of this intervention pairs well with the focus and philosophy of MCC Respectful Conversations, making it a good match. Moreover, this intervention was the fourth most effective intervention at reducing political animosity in the mega-study (like most

interventions, it did not significantly affect undemocratic practices and political violence, Voelkel et al., 2024).

In our stacked intervention, participants are asked to watch the sympathetic personal narratives videos before attending the MCC Respectful Conversation event. Our idea is that these videos can start to open people up to the idea of learning from the people they disagree with. This initial openness then primes people to have an open mind when attending the full event and, hopefully, helping to increase the effectiveness of the event. If this is successful, it would suggest that adding just a 7-minute boosting intervention can help improve the existing 150–180-minute long MCC Respectful Conversation program. Notably, this is one of the few, if only, experimental tests of possible improvements to depolarization programs.

The Current Project

The purpose of this paper is twofold. First, we aim to evaluate if the MCC Respectful Conversations program successfully reduces political animosity. This is tested with a single-group pre-test-post-test design (see Kalla & Broockman, 2022 for a similar design). We did not preregister this aspect of our analysis, so it should be seen as exploratory. Second, we aimed to experimentally test if a combined treatment of the sympathetic personal narratives video intervention and the MCC Respectful Conversation event increases people's openness to those they disagree with compared to participating in a MCC Respectful Conversation event. We use an experimental pre-test-post-test design and a series of MCC Respectful Conversation events for several different communities. We examined the effect of these interventions on a variety of outcomes representing openness and acceptance of people with different attitudes and opinions.

Method

Participants and Events

The experiment was preregistered and includes three updates with different inclusion criteria as more events became available and other events were rescheduled.

The preregistration, including all updates, is available here:

https://osf.io/fr4y3/?view_only=595eb87d613f47ffb2574b8166f23661. Data from the experiment and code to reproduce the figures and analyses in this manuscript are available here: https://osf.io/hke9w/?view_only=71362e6c532d402d86c200d80461943f.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

All participants were people who volunteered to attend one of Minnesota Council of Churches' Respectful Conversations events starting in October 2023. The date, topic, and

sample size for the events are in Table 1.² To maintain the integrity of the MCC Respectful Conversation program, demographic data was not collected. However, an informal estimate from people who attended the events describe the samples as primarily White and an even mix of men and women. We did not collect any quantitative data on the extent participants disagreed with each other. Part of the premise of the program is that the community where it is hosted has a disagreement, either within its membership or more broadly in the larger community, and people there want to address that disagreement with more structure after failing to talk about it normally. The events themselves were held for a range of communities, including municipality, school district, church, and university communities.

The original preregistration (Preregistration 1, submitted 13 October 2023) said we would collect data until the end of August 2024. Because some events were rescheduled, the preregistration was updated (Preregistration 2, submitted 15 August 2024) to extend our deadline until the 2024 November elections. An additional reschedule pushed back the event until 12 November 2024 and so we submitted a third update (Preregistration 3, submitted 13 October 2024) specifying this as our end date. During the final wave of MCC Respectful Conversation events that occurred in November 2024, there were two additional events after our 12 November 2024 cutoff for which we have data. For the analyses below, we show all results for each possible preregistration, as well as for all of the data (i.e. with the additional events after the 12 November cutoff). The events included in each preregistered inclusion criteria are noted in Table 1. Finally, one additional event was not included in the analysis because of an error in the distribution of the materials.

Design

We used a pre-post experimental design (see Figure 2 for a summary; for benefits of this design, see Clifford et al., 2021). This allowed us to test intervention stacking experimentally, as well as evaluate the intervention by comparing pre-tests and post-tests. After signing up for the event and in the 2-3 days before the event, participants received emails inviting them to complete a pre-event survey to help MCC Respectful Conversations evaluate their program and to help prepare the participant for the event. The survey included our key measures (see below), as well as questions to create anonymized codes to link participants with their post-test data.

After the questions, participants were randomly assigned to the intervention condition. This is the sympathetic personal narrative intervention from Voelkel and

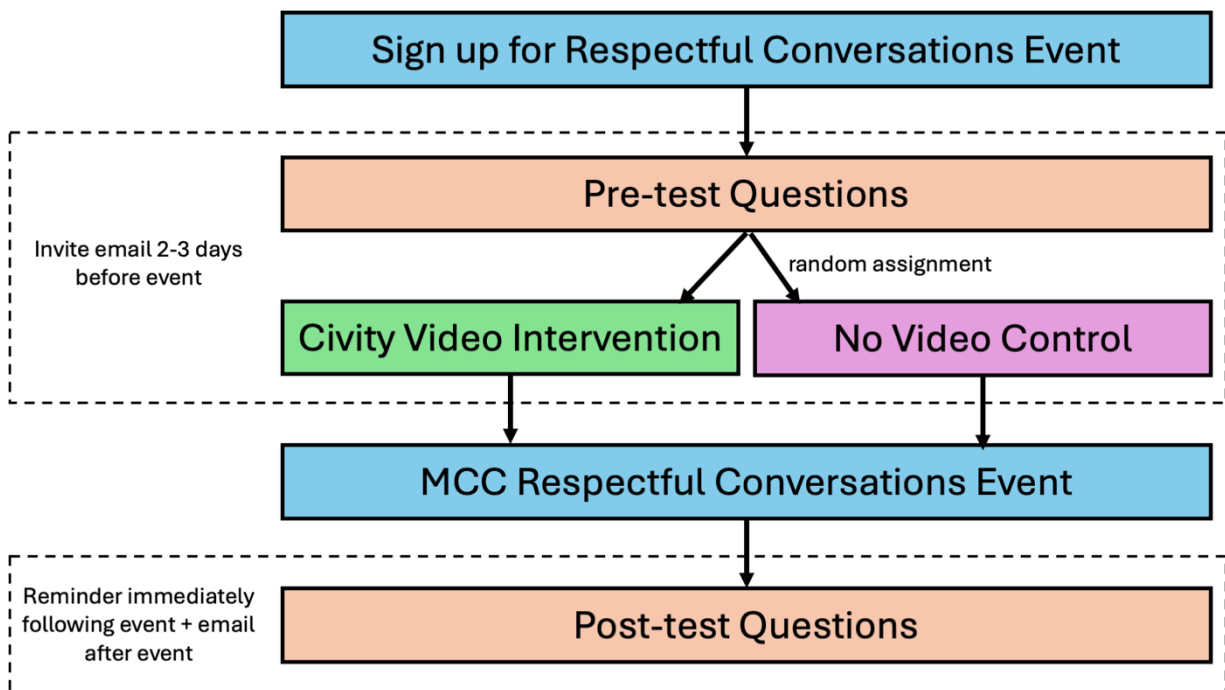
² A model including all pre-treatment measures, condition, and an event dummy code did not significantly predict attrition. We also calculated inverse probability weights following the approach of Voelkel et al. (2024) and including these weights did not change conclusions.

colleagues (Voelkel et al., 2024) developed by Civity (Kopell et al., 2024). Participants watched five short videos of people talking about what others may miss about them. Participants then watched another animated video about how democracy allows people with different views to work together. Our only modification of the original implementation is that we removed explicit mentions of politics in the text introduction to these videos because MCC Respectful Conversation events are not always about politics. We chose the Civity intervention because it was possible to implement into a MCC Respectful Conversation event and it has been shown to reduce partisan animosity (Voelkel et al., 2024). Participants randomly assigned to the control condition did not watch any videos and proceeded to the end of the questionnaire. All participants then attended their MCC Respectful Conversation event. At the event, they participated in the typical MCC Respectful Conversation process (summarized in Table 2). This design allows us to estimate the incremental, boosting change of the Civity intervention, compared to only participating in the MCC Respectful Conversations.

Table 1. *Sample sizes for each event at each time point in each condition, as well as total sample size and the median completion date of the post-test.*

Event	Date	Topic	Pre-Test Sample Sizes			Response Rate	Post-Test Sample Sizes			Mdn. Completion Date
			Control Condition	Treatment Condition	Total		Control Condition	Treatment Condition	Total	
1 ^a	2023-10-14	Bridging Cultural Differences	5	1	6	0.67	3	1	4	2023-10-15
2 ^a	2023-11-13	Co-Creating Community at [University]	23	25	48	0.90	21	22	43	2023-11-13
3 ^a	2024-03-16	Being a Reconciling Church and What That Means	10	11	21	1.00	10	11	21	2024-03-16
4 ^a	2024-08-27	This Election	15	17	32	0.97	15	16	31	2024-08-27
5 ^a	2024-08-28	This Election	13	8	21	0.81	10	7	17	2024-08-28
6 ^b	2024-10-01	This Election	10	6	16	0.75	6	6	12	2024-10-01
7 ^b	2024-10-23	This Election	40	29	69	0.75	31	21	52	2024-10-23
8 ^c	2024-11-12	The Election	7	2	9	0.78	5	2	7	2024-11-12
9	2024-11-20	The Election	15	11	26	0.58	10	5	15	2024-11-20
10	2024-11-20	The Election	8	3	11	0.64	5	2	7	2024-11-20
Total			146	113	259	0.78	116	93	209	

Note: Events sharing superscript ^a were part of Preregistration 1's inclusion criteria. Events sharing superscript ^a and ^b were part of Preregistration 2's inclusion criteria. Events sharing superscript ^a, ^b, and ^c were part of Preregistration 3's inclusion criteria. Events without any superscript were not part of any preregistrations inclusion criteria.

Figure 2. *Summary of study design***Table 2.** *Summary of MCC Respectful Conversation Event.*

Step	Activity
I	Shared meal
II	Event welcome with stated purpose, naming the framing topic
III	Video: humorous animation of physiological response to conflict Video: Illustration of guiding principles of the conversation, with improv actors
IV	humorously violating each
V	In small table groups: reminder of the guiding principles and sign of agreement
VI	In small table groups: facilitated conversation - Circle Question 1 - Circle Question 2 - Circle Question 3 - Unstructured time for participants to ask and answer questions of genuine interest - Closing Question 4
VII	Large group open reflection
VIII	Evaluation

At the end of the event, participants were encouraged to fill out the post-event questionnaire, which included our key measures and questions to create anonymized codes to link with the pre-test measures. We also sent them a reminder email the day after the event to encourage more responses. The response rates for each event and in total are in Table 1.

Measures

We had several sets of key measures that we measured at both the pre-test and post-test to assess the effectiveness of MCC Respectful Conversations and the experimental treatment (see Table 3). All measures are coded so that higher scores indicate more openness to people and ideas that differ from the participants' own point of view. All measures are rescaled to range from 0 to 1. Correlations between the scales are in the supplemental materials.

Our key measures came from two sources. First, we adopted and adapted items from the Social Cohesion Impact Measure (SCIM; Bridging Movements Goals & Measures Program, 2024), a measure designed to measure the impact of depolarization initiatives. This measure was developed in consultation with both practitioners and academics and includes multiple measures from the research literature (e.g., feeling thermometers, social distance measures). The SCIM assesses several outcomes related to polarization and we include those assessing liking of the outgroup (liking scale, items 1 and 2), morality (item 3), and respect (item 4). We also adapted items from previous evaluations of MCC Respectful Conversations that assess confidence in discussing difficult issues (items 5 and 6) and empathy (items 7 and 8).³ As preregistered, we created 4 two-item scales from these items (all between item r 's > .50, all test-retest reliability r 's $\geq$.50, see Table 3) and include item level analyses in the supplemental materials.

Table 3. *Scale items (1st column), correlations/alphas for items used in the scales (2nd and 3rd column), as well as the correlation between the pre- and post-test measurements of all scales and items (4th column).*

Scales/Items	Item r / α at Pre-test	Item r / α at Post-test	Pre-Post r
Liking Scale	.64	.68	.67
1. Please indicate how you feel toward people who disagree with you on [topic] using the scale below. 10 means that you feel very favorably or warm toward them, 0 that you feel very unfavorable or cold, and 5 is neutral.			.60

³ Prior evaluations asked participants to retrospectively report if the event increased their confidence and empathy. However, it is not clear that participants have adequate insight into whether or not their attitudes and beliefs change (Graham & Coppock, 2021).

2. How comfortable are you having friends who disagree with you on [topic]? [1 = not at all comfortable, 5 = extremely comfortable]			.61
Respect Scale	.51	.64	.68
3. Would you say that people who disagree with you on [topic] are generally good people? [1 = not at all good people, 5 = extremely good people]			.58
4. I respect the opinions of people who disagree with me on [topic]. [1 = strongly disagree, 5 = strongly agree]			.62
Confidence Scale	.88	.83	.63
5. How confident are you in discussing difficult issues with people you disagree with? [1 = not confident at all, 5 = extremely confident]			.60
6. I am confident I can talk about controversial issues with people I disagree with. [1 = strongly disagree, 5 = strongly agree]			.61
Empathy Scale	.55	.63	.63
7. I understand others' points of view on [topic]. [1 = strongly disagree, 5 = strongly agree]			.59
8. I have empathy for those whose viewpoints are different from mine on [topic]. [1 = strongly disagree, 5 = strongly agree]			.50
Moral Humility Scale	.57	.54	.62
9. I am open to learning about different moral values from people I typically disagree with. [1 = strongly disagree, 5 = strongly agree]			.44
10. My views about moral issues today may someday turn out to be wrong. [1 = strongly disagree, 5 = strongly agree]			.62
11. My moral ideas are much more just and fair than the ideas of those who disagree with me. (reverse scored) [1 = strongly disagree, 5 = strongly agree]			.52

Note: [topic] was filled in using the phrase from the topic column of Table 1.

As part of preregistered secondary analyses, we included three items from a moral humility measure (items 9, 10, and 11; Vallabha, 2025). These items come from different subscales intending to assess Moral Learning/Openness (item 9), Moral Fallibility (item 10), and Moral Superiority (item 11). We analyze these as scales in the main text and include item level analyses in the supplemental materials (α and test-retest reliability r 's in Table 3).

Results

Pre-Post Effect of MCC Respectful Conversations (Exploratory)

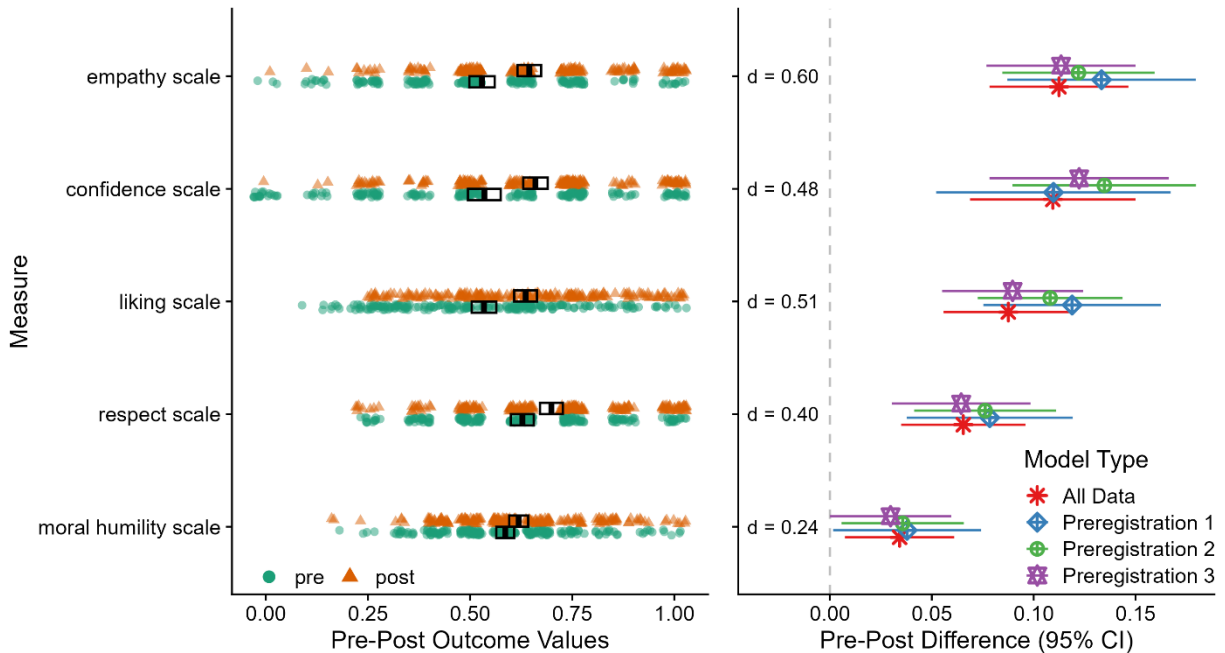
We first analyzed if there was evidence that MCC Respectful Conversation events depolarized participants. These analyses were not preregistered and so should be considered exploratory. To evaluate the trajectory of people's attitudes over the course of the events, we calculated the difference between the pre-test and post-test measures for each participant. Because participants were recruited from different events, we adjusted

for these events using OLS regression. We regressed the difference score on a series of event contrast codes using a simple coding scheme.⁴ The intercept in a simple coding scheme corresponds to the mean of event means (see UCLA: Statistical Consulting Group, 2024). The significance of the intercept tells us if there was significant change over the course of the event. We repeated this analysis for all scales and items. We also repeated it for each of our preregistered inclusion criteria, as well as for all of the data.

Across all scales, we see that people became less polarized following the MCC Respectful Conversation events (Figure 3). This was also the case for the individual items (with exceptions for two of the moral humility items, see supplemental materials). As is evident in Figure 3 (right panel), the effects were largest for the empathy and confidence scales, with both increasing by between 15% and 18% of the scale range depending on the inclusion criteria (equivalent to .60-.72 scale points on a 5-point scale). The smallest changes were for the moral humility scale with increases ranging from 6% to 7% of the scale range depending on the inclusion criteria (equivalent to .24-.28 scale points on a 5-point scale). Based on these analyses, it appears that MCC Respectful Conversation events work (see Discussion for the limitations of this single-group pre-post design).

Figure 3. *Left Panel: Outcome values (x-axis) for pre-post measures of all scales (y-axis) using the full dataset. Boxes are the means (thick line) and 95% confidence intervals (left and right edge of the box) of the scales at the pre-test and post-test for the full data set. Right Panel: Estimated pre-post difference (x-axis) for all scales (y-axis) using different subsets of data based on the preregistered inclusion criteria. Standardized effect estimates (d) are standardized by the standard deviation of the difference score using all the data.*

⁴ In response to a reviewer's question, we repeated these analyses using a multilevel model instead of the contrast codes. The conclusions are identical.



Treatment Effect of Boosting Intervention (Confirmatory)

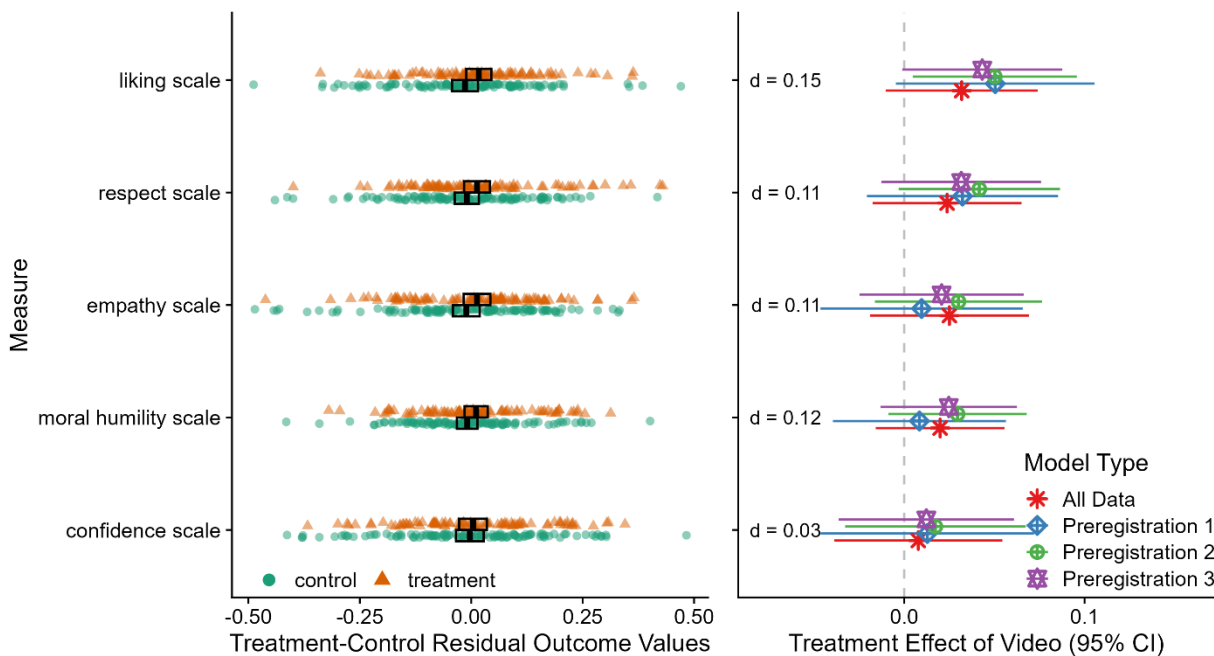
We tested our preregistered experimental hypothesis across each of the scales using OLS regression. We regressed the post-test measure on the boosting treatment condition (0 = control, 1 = treatment) and the pre-test measure to estimate the treatment effect (Clifford et al., 2021). Because participants were recruited from different events, we preregistered that we would include a fixed-effect for event as a series of dummy codes. Instead of using dummy codes, we used the simple coding scheme described above.

Across all analyses, there was one statistically significant result (Figure 4). When using the inclusion criteria from Preregistration 2, there was a significant ($p = .03$) effect of the boosting treatment, indicating that the treatment+MCC Respectful Conversations increased liking by 5% of the scale range (equivalent to .2 scale points on a 5-point scale) compared to the control condition which only included MCC Respectful Conversations. The effect on liking using the Preregistration 1 inclusion criteria was nearly identical in size, but with a larger standard error making it non-significant ($p = .07$). The boosting treatment effects using the other inclusion criteria, which all included data from after the election, were smaller in size and non-significant (p 's = .06 and .13).

There were no significant boosting effects for any of the other scales and inclusion criteria, with effect sizes that ranged from nearly zero to small and positive (see Figure 4). In addition to the liking scale reported above, the respect scale had the most consistently positive estimates across the inclusion criteria, ranging between 2% and 4% of the scale

range (equivalent to .08-.16 scale points on a 5-point scale). Although none of these effects were statistically significant (all p 's > .07).

Figure 4. *Left Panel: Residual post-test values after removing the effect of the pre-test measure and the event contrast codes (x-axis) for all scales (y-axis) using the full dataset. Boxes are the means (thick line) and 95% confidence intervals (left and right edge of the box) of the scales for the full data set. Right Panel: Estimated treatment effect (x-axis) for all scales (y-axis) using different subsets of data based on the preregistered inclusion criteria. Standardized effect estimates (d) are standardized by pretreatment standard deviation using all of the data.*



General Discussion

We evaluated the MCC Respectful Conversations program in an exploratory analysis. We also tested if using a light touch boosting intervention could improve the existing, more intensive intervention when they are simply stacked together. The idea was that the sympathetic personal narratives video intervention could help prepare participants for the event, making them more open and receptive to MCC Respectful Conversations. If successful, this would provide an easy method to boost the effectiveness of existing interventions. We tested this with a confirmatory experiment.

We found that MCC Respectful Conversation events were associated with more positive attitudes, empathy, and moral humility in a single-group pre-post design. We also found that it increased confidence in discussing difficult issues, a key skill taught in the program. Although not part of our preregistered experiment, this is useful evidence suggestive of the benefits of MCC Respectful Conversations. MCC Respectful Conversations and other interventions like it are a type of bundled intervention. That is, they bundle together multiple possible treatments that could improve outcomes. This focus on creating a program that works is different from parts of the experimental literature that focuses on isolating the effect of specific mechanisms. For example, MCC Respectful Conversations has multiple components that could impact its effectiveness. They bring people together to eat a meal (building community), teach people skills of listening and processing disagreeable information (reducing defensiveness), give people a safe space to both share and listen to personal narratives (cultivating empathy), and expose people to the genuine thoughts and feelings of people they disagree with (fostering positive contact and reducing misperceptions). Although the precise mechanism contributing to our findings is not known, experimental tests of each of these possible mechanisms are present in the literature (e.g., Hartman et al., 2022; Kalla & Broockman, 2023; Kubin et al., 2021; Lees & Cikara, 2021; Moore-Berg et al., 2020; Voelkel et al., 2021).

Notably, because our evaluation of MCC Respectful Conversations lacks a control group, to treat the estimates of this pre-post design as a causal estimate, we need to assume that MCC Respectful Conversations was responsible for the changes. There are threats to this assumption, some more plausible than others. For example, it is possible that something happened between the pre-test and post-test that caused the changes we observe (e.g., events that reduced the psychological consequences of disagreement in the USA generally); however, given the relatively short time-span between the measures and that the events were held at different times throughout the year, this does not seem plausible. Similarly, we doubt the effects are due to maturation or other developmental processes; for the adults who attend these events, such processes are unlikely to occur over the course of a few days. A greater threat to the validity of this causal estimate is that participants may report more positive attitudes, empathy, and moral humility because that is what they expect from the event. Therefore, future tests of MCC Respectful Conversations (and similar programs) should use placebo control-group designs, measures less susceptible to controlled responses, and measures of longer-term effects.

With these limitations in mind, we think that our project contributes to the growing literature evaluating practitioner-lead depolarization events. The other evaluations of in-person events found similar depolarizing effects (Baron et al., 2025; Kubin et al., 2026; see also Kalla & Broockman, 2022; Fishkin et al., 2021). Similar programs, conducted online,

show mixed results for depolarization (Blair et al., 2025a, 2025b). Future work should investigate if online events or in person events are more effective and what features of events affect effectiveness estimates (e.g., different samples attracted to online vs in person events).

Our preregistered experiment tested if Civity's sympathetic personal narrative videos (Kopell et al., 2024) could incrementally boost the effects of MCC Respectful Conversations. It primarily resulted in null estimates of incremental changes. One possibility is that the MCC Respectful Conversations events are powerful events. As discussed, they bundle together several strategies targeting multiple mechanisms relevant to depolarization. Even the participants in our control condition experienced this impactful event. Thus, the lack of an incremental effect may be because the relevant psychological mechanisms were at ceiling in the control condition, limiting the ability of the light touch intervention to cause additional changes.

This could suggest that a light-touch intervention does not boost at all (i.e. there is no effect), or it may suggest that the observed effect is smaller than anticipated (i.e. there is low statistical power). Consistent with the possibility of lower power, the values in the confidence intervals for nearly all of the scales include both zero, but also effects that would be substantively important (especially when considering the low cost of the boosting intervention).

For this reason, it is helpful to put the effect sizes of previous tests of the intervention into context with the effect sizes in our study of incremental change. The original test of the sympathetic personal narrative intervention found an effect size equivalent to $b = .09$ on a measure similar to our liking scale (Voelkel et al., 2024). Our effects are smaller than this, however, there are important differences between our study and the original test. The original test had dependent measures immediately follow the intervention and the control condition was a *neutral* control condition. That is, where our study compared the Civity video intervention to an intensive control condition, the original study compared the video intervention to much less intensive control. Moreover, in our study, the intervention was often further removed from the dependent measures.

Given this context, we might expect somewhat smaller effect sizes in our design than $b = .09$. Notably, the original test also measured their dependent measures after 2 weeks. After this delay, the effect size was equivalent to $b = .027$ or $b = .018$ (depending on sample inclusion criteria) on a measure similar to our liking scale. This is *smaller* than the effects we observed. Interestingly, the original study (Voelkel et al., 2024) also included a more active control condition. When compared to this active control, the immediate effect of the video intervention had an effect size equivalent to $b = .005$. This is *smaller* than the

effect we estimated. This is to say, that the effect sizes we observe when testing incremental changes are consistent with those from Voelkel et al (2024) when considering time delays and neutral vs active control conditions. It is an open question as to whether we would have found larger effects with more immediate measurements of the outcomes, or if our effects would last at later measurement points (e.g., weeks or months later). Generating lasting effects of depolarization initiatives appears difficult (Dias et al., 2024; Holliday et al., 2025; Voelkel et al., 2024). To initiate lasting change, it is likely necessary to also initiate changes in our communities and larger political system.

It is potentially a concern that some of the events took place so soon after a contentious election. Although we might hope that depolarization interventions, such as MCC Respectful Conversations and the sympathetic personal narrative videos, would work well in contentious environments, a major election may be too strong of a situation to overcome. Testing depolarization interventions in different political environments, with different types of political competition and outcomes will help future studies understand the extent these interventions generalize.

Finally, due to the nature of the evaluation procedure, we were unable to collect background data on the participants. If the boosting video intervention or the MCC Respectful Conversation events have different effects on people from different backgrounds, we are unable to identify that heterogeneity (our smaller sample size would also make this a challenge).

Coda

Learning how to discuss difficult topics with empathy across meaningful differences is an important skill in a pluralistic society where disagreement is inevitable. In many such cases, efforts to depolarize and maintain civility are valuable for learning others' perspectives and identifying shared interests. However, goals related to depolarization and civility are not the only relevant goals and they may not be the best goals to pursue in some situations. For example, there is initial evidence that depolarization programs might make some people less likely to confront misinformation and racism (Blair et al., 2025a). Cultivating empathy might also be less motivating than other emotions, like moral outrage, for motivating collective action and peaceful protest in responses to injustice (e.g., Spring et al., 2018; Tager et al., 2011). Rather than seeing depolarization and civility as ultimate goals, scientists and practitioners should see these goals as one set of possible goals and tools to draw on as we seek to better society.

References

- Ahler, D. J., & Sood, G. (2018). The parties in our heads: Misperceptions about party composition and their consequences. *The Journal of Politics*, 80(3), 964-981.
- Baron, H., Blair, R. A., Choi, D. D., Gamboa, L., Gottlieb, J., Robinson, A. L., ... & West, E. A. (2025). Couples therapy for a divided America: assessing the effects of reciprocal group reflection on partisan polarization. *Political Behavior*, 1-26.
- Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social issues and policy review*, 3(1), 141-177.
- Blair, R. A., Gottlieb, J., Schenk, M., & Woods, C. (2025a). Depolarization and its discontents: Experimental evidence on affective polarization and willingness to confront racism and misinformation from two online workshops. *Political Psychology*, 47(2). <https://doi.org/10.1111/pops.70046>
- Blair, R. A., Gottlieb, J., Schenk, M., & Woods, C. (2025b). Depolarizing within the Comfort of Your Party: Experimental Evidence from Online Workshops. *Political Communication*, 42(6), 925–950. <https://doi.org/10.1080/10584609.2025.2469097>
- Brandt, M. J., & Crawford, J. T. (2020). Worldview conflict and prejudice. In *Advances in experimental social psychology* (Vol. 61, pp. 1-66). Academic Press.
- Bridging Movement Goals & Measures Program (2024). *Social cohesion impact measurement*. <https://www.listenfirstproject.org/goals-and-measures>
- Byrne, D. (1969). Attitudes and attraction. In *Advances in experimental social psychology* (Vol. 4, pp. 35-89). Academic Press.
- Byrne, D., & Griffitt, W. (1966). A developmental investigation of the law of attraction. *Journal of Personality and Social Psychology*, 4(6), 699.
- Clifford, S., Sheagley, G., & Piston, S. (2021). Increasing precision without altering treatment effects: Repeated measures designs in survey experiments. *American Political Science Review*, 115(3), 1048-1065.
- Crawford, J. T., Brandt, M. J., Inbar, Y., Chambers, J. R., & Motyl, M. (2017). Social and economic ideologies differentially predict prejudice across the political spectrum, but social issues are most divisive. *Journal of personality and social psychology*, 112(3), 383.
- Dias, N. C., Aarslew, L. F., Frederiksen, K. V. S., Lelkes, Y., Pradella, L., & Westwood, S. J. (2024). Correcting misperceptions of partisan opponents is not effective at treating democratic ills. *PNAS nexus*, 3(8), pgae304.
- Gift, K., & Gift, T. (2015). Does politics influence hiring? Evidence from a randomized experiment. *Political Behavior*, 37, 653-675.

- Gower, K., Cornelius, L., Rawls, R., & Walker, B. B. (2020). Reflective structured dialogue: A qualitative thematic analysis. *Conflict Resolution Quarterly*, 37(3), 207-221.
- Graham, M. H., & Coppock, A. (2021). Asking about attitude change. *Public Opinion Quarterly*, 85(1), 28-53.
- Hartman, R., Blakey, W., Womick, J., Bail, C., Finkel, E. J., Han, H., ... & Gray, K. (2022). Interventions to reduce partisan animosity. *Nature Human Behaviour*, 6(9), 1194-1205.
- Holliday, D. E., Lelkes, Y., & Westwood, S. J. (2025). Why depolarization is hard: Evaluating attempts to decrease partisan animosity in America. *Proceedings of the National Academy of Sciences*, 122(39), e2508827122.
- Huddy, L., & Yair, O. (2021). Reducing affective polarization: Warm group relations or policy compromise? *Political Psychology*, 42(2), 291-309.
- Kalla, J. L., & Broockman, D. E. (2022). Voter outreach campaigns can reduce affective polarization among implementing political activists: Evidence from inside three campaigns. *American Political Science Review*, 116(4), 1516-1522.
- Kalla, J. L., & Broockman, D. E. (2023). Which narrative strategies durably reduce prejudice? Evidence from field and survey experiments supporting the efficacy of perspective-getting. *American Journal of Political Science*, 67(1), 185-204.
- Kubin, E., Barker, E., Keylock, L., & Gray, K. (2026). Promoting Empathy, Connection, and Pro-Sociality Through Storytelling in American High Schools: A Collaboration Between Academics and Practitioners. *PsyArxiv*. Retrieved from osf.io/preprints/psyarxiv/tghwp_v2
- Kubin, E., Puryear, C., Schein, C., & Gray, K. (2021). Personal experiences bridge moral and political divides better than facts. *Proceedings of the National Academy of Sciences*, 118(6).
- Kopell M., Strand P., Baleria G., & Fiorella M. (2024) Civity Storytelling Intervention: Expanding the Pool of People Who Matter. Civity. Retrieved October 9, 2025, from https://sshs.qualtrics.com/jfe/form/SV_88PAor84VsO3X5Y?Condition=Civity_Storytelling, also available at www.civity.org/civity-storytelling-intervention/
- Lees, J., & Cikara, M. (2021). Understanding and combating misperceived polarization. *Philosophical Transactions of the Royal Society B*, 376(1822), 20200143.
- Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization?. *The Journal of Politics*, 80(1), 59-70.
- Mernyk, J. S., Pink, S. L., Druckman, J. N., & Willer, R. (2022). Correcting inaccurate metaperceptions reduces Americans' support for partisan violence. *Proceedings of the National Academy of Sciences*, 119(16), e2116851119.

- Moore-Berg, S. L., Ankori-Karlinsky, L. O., Hameiri, B., & Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proceedings of the National Academy of Sciences*, 117(26), 14864-14872.
- Morey, J. (2013). Not changing minds but softening hearts. *Creative Nursing*, 19(4), 210-213.
- Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual review of psychology*, 72(1), 533-560.
- UCLA: Statistical Consulting Group. (2024). *R library contrast coding systems for categorical variables*. <https://stats.oarc.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/>
- Santoro, E., & Broockman, D. E. (2022). The promise and pitfalls of cross-partisan conversations for reducing affective polarization: Evidence from randomized experiments. *Science Advances*, 8(25), eabn5515.
- Shih, M., Wang, E., Trahan Bucher, A., & Stotzer, R. (2009). Perspective taking: Reducing prejudice towards general outgroups and specific individuals. *Group Processes & Intergroup Relations*, 12(5), 565-577.
- Spring, V. L., Cameron, C. D., & Cikara, M. (2018). The upside of outrage. *Trends in Cognitive Sciences*, 22(12), 1067-1069.
- Tagar, M. R., Federico, C. M., & Halperin, E. (2011). The positive effect of negative emotions in protracted conflict: The case of anger. *Journal of Experimental Social Psychology*, 47(1), 157-164.
- Vallabha, S. (2025). *Moral Humility: An Antidote to the Dark Side of Morality* (Publication No. 31933272) [Doctoral dissertation, Michigan State University]. ProQuest Dissertations & Theses Global.
- Voelkel, J. G., Ren, D., & Brandt, M. J. (2021). Inclusion reduces political prejudice. *Journal of Experimental Social Psychology*, 95, 104149.
- Voelkel, J. G., Stagnaro, M. N., Chu, J. Y., Pink, S. L., Mernyk, J. S., Redekopp, C., ... & Willer, R. (2024). Megastudy testing 25 treatments to reduce antidemocratic attitudes and partisan animosity. *Science*, 386(6719), eadh4764.
- van der Lee, R., Ellemers, N., Zarzeczna, N., & Scheepers, D. (2024). Threatened by the immoral, challenged by the incompetent: Cardiovascular responses to intragroup morality vs. competence evaluations. *Group Processes & Intergroup Relations*, 27(1), 62-75.
- Whitt, S., Yanus, A. B., McDonald, B., Graeber, J., Setzler, M., Ballingrud, G., & Kifer, M. (2021). Tribalism in America: behavioral experiments on affective polarization in the Trump era. *Journal of Experimental Political Science*, 8(3), 247-259.

Woitzel, J., & Koch, A. (2023). Ideological prejudice is stronger in ideological extremists (vs. moderates). *Group Processes & Intergroup Relations*, 26(8), 1685-1705.