Talking across the Aisle*

Luca Braghieri Peter Schwardmann Egon Tripodi

January 10, 2025

Abstract

We conduct an experiment in which U.S. Democrats and Republicans engage in naturalistic video conversations about policy-relevant facts. We investigate self-selection into politically homogeneous interactions and how these interactions affect information aggregation and affective polarization. Participants exhibit a preference against cross-partisan conversations, explained by lower expectations about their informational and hedonic value. Indeed, participants find it significantly more difficult to extract knowledge from counter-partisans and, thus, tend to learn less from them. In contrast, cross-partisan interactions prove more enjoyable than anticipated and lastingly reduce affective polarization. Overall, cross-partisan contact may better serve to reduce affective polarization than to improve information aggregation.

Keywords: cross-partisan interactions, partisan sorting, echo chambers, information diffusion, affective polarization, misperceptions.

JEL codes: C93, D83, D9.

---

* Braghieri: Bocconi University, CEPR, CESifo, and IGIER; luca.braghieri@unibocconi.it. Schwardmann: Carnegie Mellon University, CEPR, CESifo; pschwardmann@gmail.com. Tripodi: Hertie School, CESifo; egontrpd@gmail.com. We are grateful to Francesco Capozza, John Conlon, Jessica Gottlieb, Alex Imas, Ro'ee Levy, Chris Roth, Klaus Schmidt, and Mateusz Stalinski as well as seminar participants at the ESA North America Meeting, Polarize Conference in Bergen, BEDI conference in Pittsburgh, CESifo Behavioral conference, University of Bonn, Frankfurt School and Goethe University Frankfurt, University of Warwick, Ohio State University, LMU Munich, and University of Zurich for useful comments and suggestions. Yves Le Yaouanq played an integral part in the genesis of the project. The authors gratefully acknowledge funding from the "Democracy in the 21st Century" research area of CIVICA. Tripodi acknowledges funding by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119). Ethics approval was obtained at the Hertie School (approval # 20221101-07). The experiments were pre-registered in the AsPredicted registry (#155100 and #173633). We thank Eleonora Guseletova, Susanna Hasinnen, Riya Kejriwal, and Myung Won (Misha) Seong for excellent research assistance and Marlene Bargou for programming the experimental interface. All errors are our own.

1. Introduction

Over the past two decades, scholars, public intellectuals, and policymakers have raised concerns about the negative effects of political echo chambers, online and offline environments in which individuals primarily interact and share information with co-partisans rather than counter-partisans (Bishop and Cushing, 2008; Sunstein, 2001).[1] Such concerns are generally articulated along two distinct dimensions (Bishop and Cushing, 2008; Sunstein, 2001, 2017). The first dimension relates to information: echo chambers might impair the aggregation of information that is differentially distributed across party lines, thus lowering the quality of political decision-making. The second dimension relates to social cohesion: the siloing of people into homogeneous groups reduces opportunities for cross-partisan contact that might foster empathy with and understanding of individuals on the other side of the political aisle. Despite considerable public and academic interest in echo chambers and the significant hope policymakers place in initiatives that foster cross-partisan contact, there remain significant gaps in our understanding of whether and how people self-select into politically homogeneous interactions and of the consequences of these interactions for information sharing and, to a lesser extent, social cohesion.

In this paper, we make progress on these questions by engaging U.S. Democrats and Republicans in naturalistic, face-to-face video conversations about policy-relevant facts, either with co-partisans or counter-partisans. Despite the unscripted nature of the conversations, we retain substantial experimental control, enabling us to examine three key dimensions: (i) self-selection into co- vs. cross-partisan interactions, measured by participants' willingness to pay for conversations; (ii) expected and actual learning from the conversations; and (iii) the affective consequences of the conversations.

The experiment proceeds as follows. After providing their first names or aliases, political leanings, and feelings about their own and the other party, participants take an initial quiz. The quiz consists of 14 factual multiple-choice questions covering topics that, according to an exploratory survey run prior to the experiment, both Democrats and Republicans

[1] The discussion about online echo chambers was sparked by the diffusion of the internet and the rise of social media (Sunstein, 2001; Gentzkow and Shapiro, 2011); the discussion about offline echo chambers was primarily fueled by the increased degree of political homophily in patterns of geographic sorting (Bishop and Cushing, 2008; Brown et al., 2024).

deem contentious and relevant to policy (e.g., immigration, policing, healthcare), alongside questions covering basic knowledge of U.S. politics (e.g., naming the Speaker of the House). After completing the quiz, participants are informed of the opportunity to have a conversation with another participant in the study. Depending on the treatment assignment, the conversation partner is either a co-partisan or a counter-partisan. Participants are informed of the political affiliation of their conversation partner and of the fact that, after the conversation, they will be given a chance to revise their answers to the quiz.[2] Participants then state, in an incentive-compatible elicitation, i) their willingness to pay for having the conversation, and ii) their expected improvement in the number of correct answers to the quiz as a result of having the conversation. The willingness-to-pay decision is implemented for 5 percent of the sample. Most of our analyses focus on the remaining 95 percent of the sample, which, independently of their willingness to pay, is released into an eight-minute video conversation with a fellow participant. After the conversation, participants are given the opportunity to revise their answers to the quiz. Participants are then asked to predict their expected improvement one more time before answering a series of questions about their experience of the conversation, their feelings about their own and the other party, and their demographics.[3]

Our first finding is that participants have a higher willingness to pay for interacting with co-partisans than for interacting with counter-partisans. This relative preference for co-partisan conversations implies that partisans, if given the choice, would self-select into echo chambers. We also elicit open-ended responses about the considerations driving participants' stated preferences. Participants assigned to cross-partisan conversations are less likely to mention a desire to improve their quiz performance and more likely to mention concerns about the conversation being unpleasant. Both of these factors are highly correlated with the elicited willingness to pay and are among the most frequently stated open-ended considerations. These findings suggest that both instrumental motives (i.e., a desire to learn) and hedonic motives (i.e., the expected discomfort from an unpleasant social interaction) drive participants' relative preference for co- versus cross-partisan interactions.

[2]Whether the initial or revised answers to the quiz are payoff-relevant is resolved via a virtual coin flip and communicated to participants at the end of the experiment.

[3]We recruited participants online, through Prolific and CloudResearch, enabling us to achieve a scale and a level of demographic and geographic diversity that would be hard to achieve in an offline setting.

Our second set of results, and the main contribution of this paper, examines learning and the instrumental value of co- and cross-partisan conversations. For the first time, we are able to study both actual and expected learning from those conversations by measuring participants' knowledge before and after the interaction and comparing their beliefs to an objective ground truth. We find that participants expect co-partisan conversations to lead to significantly larger improvements in the revised quiz than cross-partisan conversations. These expectations are qualitatively correct, as co-partisan conversations actually do lead to larger improvements in the revised quiz, albeit only at marginal statistical significance ($p = 0.064$).

We can decompose this treatment gap in actual improvement into two components: potential improvement and difficulties in knowledge extraction. We show that potential improvement, defined as the number of questions that a participant does not have the correct answer to and her conversation partner does, is greater in cross-partisan conversations. This is a consequence of knowledge being distributed across party lines. At the same time, the rate of knowledge extraction, defined as actual improvement conditional on the potential to improve from the conversation, is significantly lower in cross-partisan interactions. We present evidence suggesting that greater difficulties in knowledge extraction in cross-partisan interactions are driven by participants' relative pessimism about the credibility of counter-partisans as information sources. This is consistent with a mechanism, similar in spirit to the one emphasized by Gentzkow and Shapiro (2006), whereby partisans perceive information sources that they disagree with as less credible and trustworthy.[4] To test for and ultimately rule out other explanations for the observed gap in knowledge extraction, we are able to leverage a rich dataset on the structure, content, and tone of the conversations, as well as the personal characteristics of participants.

Our third set of results focuses on the hedonic aspects of the conversations. Although participants initially expressed more concern about how much they would enjoy cross-partisan conversations compared to co-partisan ones, we find that, on average, they rate

[4]The finding that participants are overly pessimistic about the credibility of counter-partisans as information sources qualifies our overall conclusions about learning expectations. Although participants' average beliefs about the relative extent to which they can learn from co- and cross-partisan interactions are broadly correct, they seem to be based on a mistaken mental model that is too pessimistic about the other side's knowledge and, hence, the potential to learn from conversations across the aisle.

both types of interactions as equally enjoyable ex-post. Moreover, we find that cross-partisan conversations lead to a significant reduction in affective polarization, with the effect persisting in an obfuscated follow-up survey administered more than three months after the end of the experiment. The degree of affective polarization reduction after cross-partisan conversations is predicted by some personal characteristics of the participant and their conversation partner (e.g., being female, scoring lower on a social potency scale, and being randomly matched with an older partner) and by how positively they experienced the conversation.

Our findings help contextualize the widespread concerns about echo chambers as well as some of the policies proposed to address them. On the social cohesion front, we find that fact-based cross-partisan political conversations reduce affective polarization in the medium term, as conversations in other settings, often not exclusively focused on politics, have been shown to do in the short-term (Santoro and Broockman, 2022; Blattner and Koenen, 2023; Fang et al., 2023; Hobolt et al., 2024). On the instrumental front, we paint a novel and more pessimistic picture. Our results show that, even in settings where knowledge is distributed across the aisle, it might be harder for partisans to harness and share such knowledge when talking to counter-partisans. Thus, the simple policy of encouraging cross-partisan interactions might increase social cohesion but fail to significantly enhance information aggregation or affect people's propensity to sort into echo chambers for instrumental reasons going forward. Our results suggest that pessimistic beliefs about the potential to learn from counter-partisans are both a driver of self-selection away from cross-partisan conversations and a barrier to actually learning from them. As a result, policies that successfully reduce biases about the other side's informedness may lead to meaningful increases in information sharing.

This paper contributes to various strands of the literature. Motivated by the concern that echo chambers harm information aggregation and exacerbate affective polarization (Sunstein, 2001, 2009, 2017), the first of these strands establishes the existence of echo chambers, both online and offline (Braghieri et al., 2024; Brown et al., 2024; Flaxman et al., 2016; Gentzkow and Shapiro, 2011; González-Bailón et al., 2023; Guess et al., 2018; Guess, 2021; Nelson and Webster, 2017). Our experiment documents the kind of self-selection that can explain the emergence of echo chambers and proceeds to isolate hedonic and instrumental motives as plausible drivers of such self-selection.

Our paper's main contribution is the investigation of the expected and actual instrumental value of cross-partisan conversations, which connects to a large literature on social learning encompassing both theoretical work (DeGroot, 1974; Banerjee, 1992; Bikhchandani et al., 1992; Morris, 2001; Jackson and Yariv, 2007; Golub and Jackson, 2010; Golub and Sadler, 2016) and empirical analyses (Banerjee et al., 2013, 2019; Barrera et al., 2020; Braghieri, 2024; Chandrasekhar et al., 2022; Conlon et al., 2021; Guriev et al., 2023; Henry et al., 2022; Fehr et al., 2024; Graeber et al., 2024). Departing from much of this literature, our experiment explores bi-directional learning through conversations — a complex and yet fundamentally human mode of engagement and information exchange. Additionally, our study uniquely examines social learning through the lens of partisan identity, making ours the first investigation of: i) the beliefs that drive partisans to favor co-partisans over counter-partisan conversations, ii) the capacity of partisans to learn from counter-partisans in political conversations, and iii) the mechanisms by which, in these conversations, partisans glean less knowledge from counter-partisans than co-partisans.[5]

Finally, our paper makes an additional, comparatively more modest, contribution to a fast-growing literature on the potential of intergroup contact to reduce affective polarization and prejudice more generally (Allport, 1954; Bazzi et al., 2019; Boisjoly et al., 2006; Corno et al., 2022; Dahl et al., 2018; Dustmann et al., 2019; Enos, 2014; Fang et al., 2023; Blattner and Koenen, 2023; Lowe, 2021; Mousa, 2020; Paluck et al., 2019; Pettigrew and Tropp, 2006; Rao, 2019; Rossiter, 2023; Rossiter and Carlson, 2024; Santoro and Broockman, 2022; Schindler and Westcott, 2021; Scacco and Warren, 2018).[6] Our paper is unique in studying factual learning alongside affect and, as we detail in a meta-analysis in Appendix G, stands

[5]A related literature, primarily employing information provision experiments, examines how individuals select and update from information sources with differing partisan labels or affiliations (Acemoglu et al., 2024; Bauer et al., 2023; Belot and Briscese, 2022; Chopra et al., 2024; Garcia-Hombrados et al., 2024; Kashner and Stalinski, 2024; Jo, 2017; Robbett et al., 2023; Burnitt et al., 2024). We build on this literature by focusing on face-to-face conversations, which: i) represent an extremely common way in which people learn in the real world, and ii) possess unique qualities that distinguish them from information provision experiments. For example, information exchange in conversations can break down when interactions become heated and confrontational. Similarly, the dynamic nature of conversations enables trust to evolve as the dialogue progresses.

[6]Alongside social segregation or lack of contact, theoretical and empirical contributions to political economics point to voter overconfidence (Ortoleva and Snowberg, 2015), competition in news provision (Perego and Yuksel, 2022), and the evolution of learning technologies (Yuksel, 2022) as important sources of increasing ideological disagreement.

out in terms of sample size, topic of discussion, and the ability to observe persistent effects on affective polarization. Santoro and Broockman (2022) find that cross-partisan conversations can reduce affective polarization in the short term, but only when participants are instructed to have conversations about a topic that is unrelated to areas of ideological disagreement. In a large German sample, Blattner and Koenen (2023) and Fang et al. (2023) show that wide-ranging and unconstrained cross-partisan conversations that feature both political and personal topics also decrease affective polarization.[7] Our findings raise the intriguing possibility that cross-partisan conversations can persistently reduce affective polarization even when they are centered on politics, as long as such conversations focus primarily on areas of factual rather than ideological disagreement.

The rest of the paper is organized as follows. Section 2 describes the experimental design and sample. Section 3 presents results, focusing in turn on treatment differences in self-selection into co- and cross-partisan conversations (Section 3.1), the informational value of conversations (Section 3.2), and the hedonic consequences of conversations (Section 3.3). Section 4 concludes.

2. Experimental Design and Sample

2.1. Structure of the Experiment

The overarching aim of the experiment is to facilitate naturalistic fact-based face-to-face conversations about politics between Democrats and Republicans, while imposing enough structure to measure i) the hedonic and instrumental motives that drive partisan sorting into political conversations, ii) the quality of information sharing in these conversations, and iii) the consequences of cross-partisan political conversations for social cohesion. The structure

[7] The samples in Blattner and Koenen (2023) and Fang et al. (2023) consist of individuals who signed up for, and thus self-selected into, cross-partisan conversations. This raises the question of whether the impact of contact on affect is limited to those already motivated to engage in such interactions. As discussed in Section 3.3.5, our experimental setting allows us to address and assuage this concern by testing for potential selection effects. Specifically: i) at recruitment, participants are not aware that the experiment might feature a cross-partisan interaction, and ii) we find no evidence that participants with a higher willingness-to-accept to engage in a cross-partisan interaction are more likely to complete the experiment or experience larger reductions in affective polarization.

of the experiment is summarized in Figure 1. Appendix I contains the full experimental instructions.

After stating their first names or aliases, political leanings, and the warmth they feel toward their own and the other party, participants are given ten minutes to complete an initial quiz consisting of fourteen factual multiple-choice questions about politics. The quiz is reported in its entirety in Appendix Table A.1. The quiz questions focus on basic knowledge of U.S. politics (e.g., naming the Speaker of the House) and on topics that, according to an exploratory survey run prior to the experiment, both Democrats and Republicans deem important (e.g., immigration, policing, healthcare). In order to incentivize the quiz, we provide a piece rate of 0.6 dollars per correct answer. For each quiz question, we also elicit participants' unincentivized confidence in their own answers and their beliefs about the fraction of Democrats and Republicans from a pilot version of the study who answered the question correctly.

After completing the quiz, participants are told that they will have the opportunity to engage in an eight-minute conversation about the quiz with another participant in the study and that, afterward, they will be given a chance to revise their answers to the quiz in light of the information obtained in the conversation. Participants are also informed that, at the end of the experiment, one of the two quizzes will be selected at random for payment. Depending on treatment assignment, the conversation partner is either a Democrat or a Republican, thus generating random variation in whether the conversation is between co-partisans or counter-partisans.[8] We introduce participants' prospective conversation partners by their first names or aliases and political affiliations. Thus, participants are informed of the political leanings of their conversation partner before entering the conversation, capturing real-world situations where political affiliations are either known or readily inferred.[9]

Next, we measure participants' incentivized expectation of how many more questions they will answer correctly in the Revised Quiz compared to the Initial Quiz. We then inform

[8]We note that 20 percent of participants drop out after treatment assignment for reasons primarily due to technical issues with the video call. Reassuringly for internal validity, Appendix Table A.2 shows no differential attrition across treatments or partisan affiliation composition.

[9]Although the political affiliation of conversation partners is often not explicitly stated, research on mediated group discussions suggests that when politically charged topics arise, most participants quickly infer each other's political affiliations (Hobolt et al., 2024)

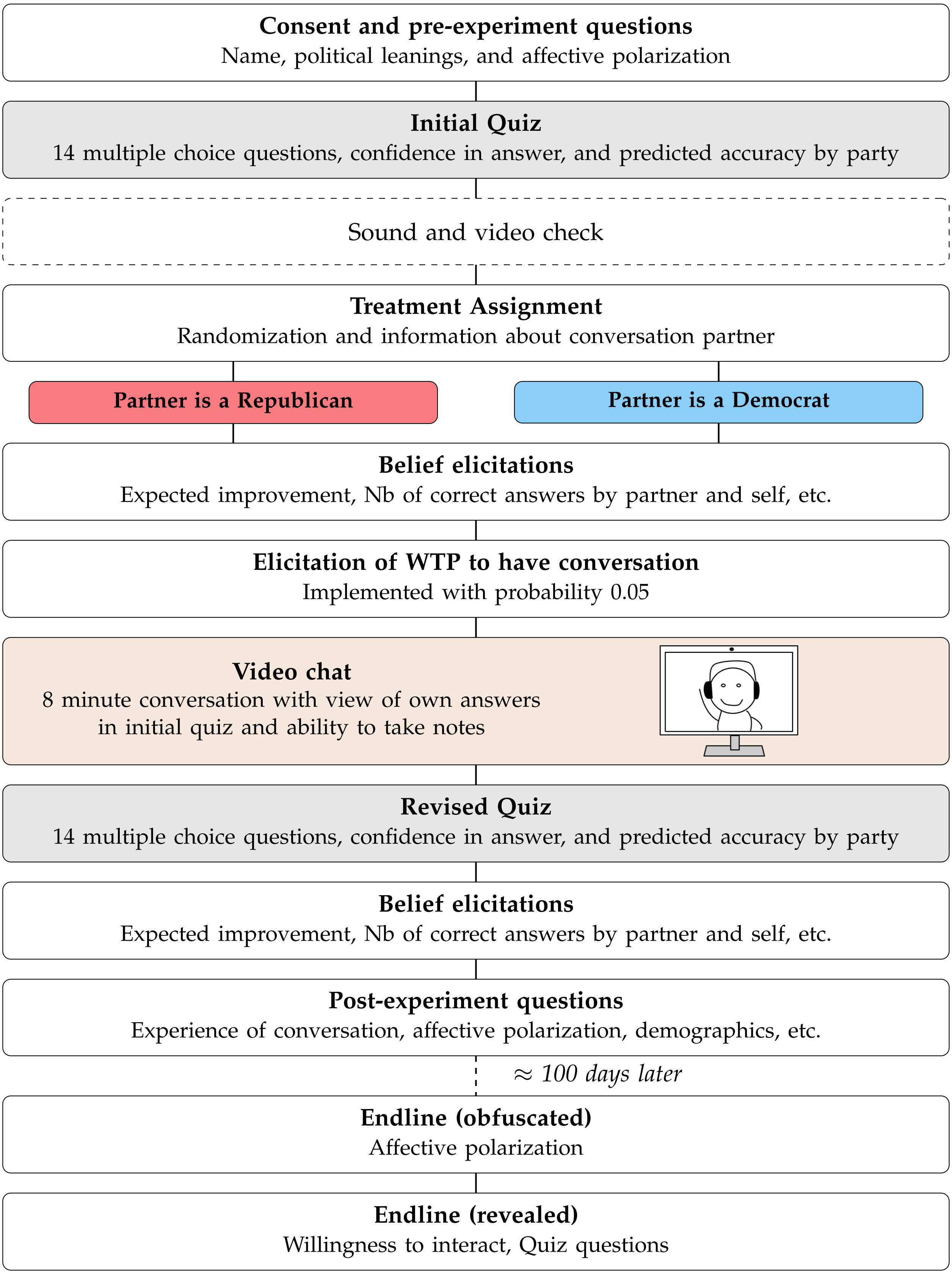

Figure 1: Design Overview

participants that, with a 5 percent probability, they will be assigned to the role of "deciders". Unlike other participants, "deciders" are given a choice as to whether or not they want to engage in the conversation with their partner. Introducing the role of "decider" allows us to elicit, from each participant, a willingness to pay to have (or avoid having) the eight-minute conversation.

All participants who are not randomly assigned to the role of deciders and who pass an audio and video check are released into an eight-minute unstructured video conversation with their partners. On one side of the screen, participants see a box showing a live video of their conversation partner as in a standard video call. On the other side of the screen, participants see their own answers to the Initial Quiz and a text box that they can use to take notes. We tell participants that they can use the video chat to discuss their quiz answers with their partner, but, other than that, the conversations are completely unstructured.

Immediately after the conversation, participants are given the opportunity to revise their answers to the quiz. While taking the Revised Quiz, participants are shown both their answers to the Initial Quiz and the notes that they took during the conversation with their partner.

After participants complete the Revised Quiz, we collect several additional measures. First, we again ask participants to report their expected improvement on the quiz as a result of having had the conversation. Second, we elicit five commonly used measures of affective polarization that we describe in more detail in the next section. Third, we elicit participants' beliefs about the extent to which they found their conversation partner knowledgeable. Fourth, we ask participants about their experience in the conversation. Finally, we elicit demographic characteristics.

Approximately 100 days after the intervention, we re-contacted our study participants for an "obfuscated" follow-up survey (Haaland and Roth, 2020). This survey had two main purposes: first, to measure the persistence of our affective polarization results and, second, to probe the sensitivity of those results to experimenter demand effects. The follow-up survey was "obfuscated" in the sense that we designed the survey in such a way as to make it seem unrelated to the main part of the experiment described above. Specifically, we modified the recruitment template, we changed the account from which we invited participants to take part in the study, we formatted the survey differently, and, at least for the first part of the survey, we made no reference to the main experiment. Only after participants answered a battery of questions about affective polarization were they informed that the survey was connected to our main study. We then administered our quiz one last time.[10] The full set of instructions for the follow-up survey can be found in Appendix J.

Discussion of design trade-offs. Our experimental design strives to strike a balance between ecological validity and experimental control. Specifically, we sought to make the conversations as naturalistic as possible to mirror the most common way in which humans typically exchange information. Allowing for naturalistic conversations is particularly important in our setting, because we wanted to capture the possibility of breakdowns in information transmission due to conversations becoming heated, confrontational, or otherwise uncooperative.

As a complement to this naturalistic approach, we introduced sufficient experimental control to accurately measure information transmission and learning. For this purpose, we incorporated a structured, albeit somewhat artificial, quiz to capture participants' best guesses about factual statements whose accuracy we could objectively verify and incentivize. To illustrate the value of this design feature, we note that some less structured experiments document a greater convergence of partisan opinions after cross-partisan conversations (Hobolt et al., 2024; Fang et al., 2023). It is tempting to interpret such convergence in opinions as a marker of learning. However, convergence in opinion and learning are distinct phenomena, and a key advantage of our design's ability to measure learning is that we can make that distinction. In particular, our experiment, in line with the existing literature, documents markers of converging opinions: participants' responses to the quiz questions become less stereotypical of their party's typical answers after cross-partisan conversations (see Section 3.2.5), and participants in those conversations increase the extent to which they value the ideas of the opposing party (see Section 3.3.2). However, these markers of convergence in opinions do not imply learning. In fact, the experimental results allow us to rule out even small positive effects of cross-partisan conversations on actual learning (see Section 3.2).

[10]A unique aspect of our follow-up survey is that the degree of obfuscation decreases as respondents progress in the survey. As a result, our primary outcome of interest — a common measure of affective polarization — is elicited under full obfuscation, whereas other outcomes are elicited under less obfuscation.

	Pre-Interaction	Post-Interaction	Several months later
Willingness to Pay	✓		
Expected Improvement	✓	✓	
Factual Knowledge	✓	✓	✓
Affective Polarization *(Thermometer)*	✓	✓	✓
Affective Polarization *(Other Measures)*		✓	✓

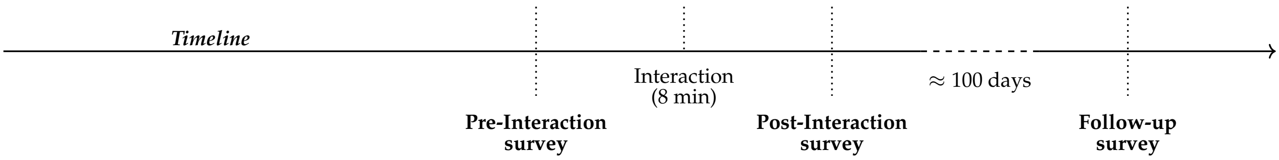

Figure 2: Timeline of Outcomes

Another feature of our experimental design warrants further discussion. By incentivizing quiz performance, our experiment artificially increases the instrumental value of the conversations and focuses them on the topics covered by the quiz. While this design choice limits our ability to assess the "natural" magnitude of the instrumental value of co- and cross-partisan political conversations in real-world settings such as elections, it enables us to explore both perceived and actual differences in the instrumental value of co- and cross-partisan interactions, as well as how instrumental considerations influence self-selection into co- vs. cross-partisan interactions. We focus on the difference in the instrumental value of co- and cross-partisan conversations, because the relative value of instrumental and hedonic motives for engaging in partisan interactions varies naturally across contexts, whereas the differences we document are likely to remain relevant in any settings where the instrumental value of politically charged conversations is high.

2.2. Main outcomes

We measure several of our main outcomes at different points in time. Figure 2 summarizes when each outcome is measured.

Willingness to pay. We use the incentive-compatible Becker-DeGroot-Marschak mechanism to measure willingness to pay to have (or avoid having) the conversation (Becker et al., 1964). Specifically, our willingness to pay elicitation proceeds in three steps. First, we ask participants whether they would be willing to give up some of their baseline payment in order to have the conversation with their partner, whether they would need to receive an additional bonus on top of their baseline payment in order to have the conversation with their partner, or whether they are indifferent between having and not having the conversation (which involves them keeping their baseline payment and receiving no additional bonus). Second, we elicit the minimum amount of money — on a scale from zero to twice the baseline payment — that would make participants indifferent between having the conversation and not having it. If participants choose a number below their baseline payment, it means they would be willing to give up some of their baseline payment in order to have the conversation; conversely, if participants choose a number above their baseline payment, it means they would need to be paid a bonus on top of their baseline payment in order to have the conversation. Third, we inform participants via a pop-up banner of any inconsistency between their categorical response in step one and their stated willingness to pay in step two. Specifically, whenever an inconsistency is detected, the pop-up banner explains the inconsistency and discourages participants from continuing to the next page. The pop-up banner disappears as soon as the inconsistency is resolved. For ease of interpretation, we subtract participants' willingness to pay from the baseline payment. This way, a positive willingness-to-pay indicates that participants are happy to pay some money in order to have a conversation with their partner, a negative willingness-to-pay indicates that participants need to be paid some money in order to have a conversation with their partner, and a willingness-to-pay of zero indicates that participants are indifferent between having and not having the conversation.

Expected improvement. Participants are asked to report their expectation of the additional number of correct answers in the Revised Quiz, which occurs after the conversation, compared to the Initial Quiz. Responses are numerical and participants are allowed to include one decimal point. To incentivize this question, we use a binarized scoring rule (Hossain and Okui, 2013) that is incentive-compatible irrespective of a participant's degree of risk aversion. The rule works as follows: the closer a participant's answer is to the realized state of the world (her actual improvement), the higher the probability that she wins a fixed bonus payment.[11] Of course, incentivizing the expected improvement question might affect

[11] Danz et al. (2022) show that simplifying the instructions of the binarized scoring rule improves the accu-

participants' behaviors in the Revised Quiz. Specifically, one might worry that participants tailor their answers in the Revised Quiz to match their stated expectations. In order to check whether incentives affect participants' answers to the expected improvement question, we incentivize the question only for a randomly drawn half of participants. As shown in Appendix Table A.3, we find no evidence that incentives affect participants' answers to the expectation question and behaviors in the Revised Quiz.

Factual knowledge. We measure factual knowledge as the score on our fourteen-question political quiz. To make sure that correct answers are strongly correlated with actual knowledge of a topic, each multiple-choice question has five possible answers, only one of which is correct. The political quiz is incentivized, but we note that there are pros and cons to incentivizing it. The obvious advantage of incentivizing the knowledge quiz is that participants are more likely to care about choosing the correct answer. The main disadvantage is that participants might attempt to cheat by, for instance, looking up the answers to the questions on the Internet.

We implemented three strategies to minimize or measure cheating behavior. First, we informed participants from the start that the study aimed to assess only what they knew or learned during the experiment, and that any outside research might be penalized. Second, we designed several of the quiz questions to be difficult to answer through online searches. Third, we introduced a surprise three-question bonus quiz near the end of the study, containing questions that could be easily answered through online searches. The bonus quiz had the same incentive structure as the other quizzes. By examining the percentage of participants who answered all three questions in the bonus quiz correctly, we can arguably gain an estimate of the number of participants who might have cheated by searching for answers online. Less than 1.8 percent of participants answered all three questions correctly, alleviating concerns about cheating.[12] Lastly, unless cheating varies significantly by treatment status, it is unlikely to affect most of our findings. The list of questions for both the main and surprise quizzes can be found in Appendix Table A.1.

Affective polarization. We include five standard measures of affective polarization (Levy,

racy of belief elicitation. Based on this insight, we provide a non-quantitative explanation of the incentives in the survey instructions and we include a link to the quantitative details for further reference.

[12]The main results remain qualitatively identical when these participants are excluded from the analysis.

2021). First, as our main measure, we employ a *feeling thermometer* by eliciting respondents' feelings towards each party's affiliates on a scale from 0 (extremely negative) to 100 (extremely positive), and we construct a measure of affective polarization based on the difference between the two. Second, we measure the *difficulty in understanding the perspectives of others*: respondents rate the perceived difficulty of understanding each party's point of view on a 5-point scale from 1 (not at all difficult) to 5 (extremely difficult). Third, we elicit the perceived *importance of considering each party's perspective* on a 5-point scale from 1 (not important at all) to 5 (extremely important). Fourth, we ask respondents to report their *perceived number of good ideas* for each party using a 4-point scale, with options ranging from 0 (almost no good ideas) to 3 (a lot of good ideas). For each of the first four measures, we compute the difference between the values elicited for one's own and the other party, re-orienting variables in such a way that larger values indicate higher levels of polarization. Fifth, we elicit the emotional reactions to hypothetical *marriages of one's own children to an out-party member*, on a 3-point scale from 0 (not upset at all) to 2 (very upset). This response is used directly, without computing a gap between attitudes towards one's own and other party members. We then derive a composite index of affective polarization by: i) standardizing each outcome, ii) orienting each outcome in such a way that higher numbers always indicate a higher degree of affective polarization, and iii) taking an equally weighted average of the standardized and re-oriented variables.

2.3. Procedures

Recruitment. We recruited study participants from both Prolific and CloudResearch Connect. The algorithm that randomly matches participants with a Democrat or with a Republican required that a sufficient number of participants take the experiment at the same time. We induced this required thickness by rolling out various recruitment surveys on a daily basis that invited participants for the main experimental session later in the day. We also used the recruitment survey to screen out Independents. Overall, we conducted 24 sessions of the experiment between December 13^{th}, 2023 and March 25^{th}, 2024.

On May 4^{th} 2024, on average 98 days after participants took the first survey, we targeted those participants again and invited them to the obfuscated follow-up study. As discussed, we used different Prolific and CloudResearch accounts to recruit participants, as well as a

Table 1: Sample demographics

	(1) Overall	(2) Democrats	(3) Republicans
Age	42.00	40.24	44.44
	(0.43)	(0.54)	(0.69)
Female	0.47	0.51	0.42
	(0.02)	(0.02)	(0.02)
White	0.77	0.71	0.85
	(0.01)	(0.02)	(0.02)
Black	0.15	0.19	0.09
	(0.01)	(0.02)	(0.01)
Asian	0.11	0.13	0.07
	(0.01)	(0.01)	(0.01)
Latino Identity	0.07	0.08	0.06
	(0.01)	(0.01)	(0.01)
Graduated College	0.22	0.25	0.17
	(0.01)	(0.02)	(0.02)
Household Income over 50k	0.70	0.67	0.75
	(0.01)	(0.02)	(0.02)
Urban Residence	0.54	0.60	0.45
	(0.02)	(0.02)	(0.02)
Voted for Trump	0.34	0.02	0.79
	(0.02)	(0.01)	(0.02)
Voted for Biden	0.55	0.87	0.09
	(0.02)	(0.01)	(0.01)
Observations	993	577	416

Notes: This table presents summary statistics for our main sample. Column (1) shows results for the overall sample, Columns (2) and (3) split the sample along party lines. Standard errors in parentheses.

different consent form.

Software. We ran the recruitment and follow-up surveys using Qualtrics. The main experiment is programmed using oTree (Chen et al., 2016), with a custom integration to the Daily API that allows us to create pair-specific video-call rooms and store the recordings of these calls.

Preregistration. The main hypotheses, experimental design, and sample size criteria for both the main experiment and the follow-up survey were pre-registered on `AsPredicted.org` (with pre-registration IDs #155100 and #173633, respectively). Pre-registration files can be found in Appendix K.

2.4. Sample and Reweighting

As shown in Table A.2, attrition was modest and not differential by treatment.[13]

[13]Following Lin et al. (2016), we also estimate a model in which our main indicator for attrition is regressed on the treatment indicator, covariates (age, gender, partisan affiliation, and baseline polarization) and the interaction of treatment indicators with covariates. An F-test of joint significance of these interaction terms allows us to test for treatment-by-covariate differences in attrition, for which we find no significant evidence ($p = 0.185$).

Table 2: Balance

	(1) Co	(2) Cross	(3) p-value (2)-(3)	(4) Co (weight)	(5) Cross (weight)	(6) p-value (5)-(6)
Age	41.491	42.539	0.226	42.110	42.539	0.631
	(0.593)	(0.630)		(0.632)	(0.630)	
Female	0.477	0.461	0.599	0.474	0.461	0.685
	(0.022)	(0.023)		(0.023)	(0.023)	
White	0.754	0.783	0.285	0.780	0.783	0.915
	(0.019)	(0.019)		(0.019)	(0.019)	
Black	0.167	0.124	0.054	0.149	0.124	0.250
	(0.017)	(0.015)		(0.016)	(0.015)	
Asian	0.098	0.114	0.431	0.090	0.114	0.217
	(0.013)	(0.014)		(0.013)	(0.014)	
Latino Identity	0.077	0.066	0.521	0.075	0.066	0.605
	(0.012)	(0.011)		(0.012)	(0.011)	
Graduated College	0.220	0.211	0.722	0.206	0.211	0.848
	(0.018)	(0.019)		(0.018)	(0.019)	
Household Income over 50k	0.695	0.715	0.503	0.704	0.715	0.713
	(0.020)	(0.021)		(0.021)	(0.021)	
Urban Residence	0.560	0.519	0.192	0.537	0.519	0.576
	(0.022)	(0.023)		(0.023)	(0.023)	
Republican	0.344	0.498	<0.001	0.500	0.498	0.949
	(0.021)	(0.023)		(0.023)	(0.023)	
Voted for Trump	0.287	0.401	<0.001	0.410	0.401	0.772
	(0.020)	(0.022)		(0.024)	(0.022)	
Voted for Biden	0.597	0.492	0.001	0.474	0.492	0.587
	(0.022)	(0.023)		(0.023)	(0.023)	
Affective Polarization (baseline)	41.552	39.897	0.361	38.989	39.897	0.626
	(1.252)	(1.310)		(1.321)	(1.310)	
Confidence in Initial Quiz	64.129	63.847	0.772	64.649	63.847	0.428
	(0.665)	(0.710)		(0.720)	(0.710)	
Score in Initial Quiz	6.487	6.568	0.668	6.446	6.568	0.527
	(0.129)	(0.138)		(0.135)	(0.138)	
Observations	509	484	993	509	484	993

Notes: This table presents a balance test for participants in our main sample. In Columns 4-6, Republican/Democrat-only pairs are weighted to address the lower number of Republican-only pairs in the sample. Robust standard errors in parentheses.

Table 1 presents summary statistics for our main sample. Consistent with known difficulties in recruiting Republicans for online experiments, our sample features around 28 percent fewer Republicans than Democrats (Kashner and Stalinski, 2024). Compared to Democrats, Republicans in our sample are slightly older, more likely to be males, more likely to be white, less likely to have a college degree, and more likely to have a household income above 50 thousand dollars.

Table 2 presents a balance table that compares the characteristics of individuals in the co-partisan and the cross-partisan interaction groups.[14] As shown in the table, the co-partisan group features fewer self-identified Republicans, fewer self-declared Trump voters, and more self-declared Biden voters. This imbalance is mechanical and stems from the interaction of two forces: i) we recruited fewer self-identified Republicans than Democrats as discussed above, and ii) the cross-partisan interaction condition necessarily features an

[14] Appendix Table A.4 provides balance tests at different stages of the experiment.

equal number of Republicans and Democrats. As a result of these two forces, there are relatively fewer Republicans than Democrats in the co-partisan group, thus mechanically generating the imbalance shown in the balance table. We address this imbalance by reweighting observations so as to mimic having an equal number of Democrats and Republicans both in the co-partisan and the cross-partisan interaction groups. We do this for all of our analyses below. Reassuringly, this procedure brings balance also to the individual characteristics that were not directly targeted, such as Black ethnicity and voting behavior, as can be seen from columns 4 through 6 of Table 2.[15]

We note that the imbalance above did not arise from a failure of randomization. Randomly matching individuals coming from a population consisting of two groups of different sizes into pairs is expected to yield such an imbalance. Despite not arising from a failure of randomization, the imbalance would bias the interpretation of the results in the absence of reweighting. Specifically, in the absence of reweighting, the comparison between cross-partisan and co-partisan pairs would disproportionately reflect the comparison between cross-partisan and Democrat-Democrat pairs. For robustness, Appendix Table A.5 reports our main analyses with an alternative (inverse probability) weighting procedure, as well as with the addition of controls.

3. Results

The results section is organized as follows. In Section 3.1, we leverage our willingness to pay measure to document a preference for self-selecting into echo chambers. We then show that this self-selection is plausibly driven by two main factors: beliefs about the instrumental value of the conversation and beliefs about the hedonic value of the conversation. In Section 3.2, we explore the instrumental value of the interaction by investigating the expected and actual informational consequences of having the conversation. Here, we first decompose the actual learning done by participants into potential for learning and knowledge extraction, and then document and explain a treatment gap in the latter component. Section 3.3 examines the hedonic effects of cross-partisan conversations: it documents an

[15]The reason why our reweighting procedure addresses the small imbalance on Black ethnicity is arguably because Black ethnicity is highly correlated with self-identifying as a Democrat.

immediate and lasting reduction in affective polarization, and then relates that finding to the broader literature on the subject.

Following our preregistration, our analyses primarily compare cross-partisan with co-partisan interactions. In Appendix E, we revisit our main results looking separately at Republicans and Democrats. The results from the split sample show a pattern similar to the aggregate comparisons.

3.1. Self-selection into Echo Chambers

We begin our analysis by studying participants' willingness to pay for co- and cross-partisan conversations, which we interpret as a holistic measure of participants' preferences for interacting with co- and counter-partisans.

3.1.1. Treatment Differences in Participants' Willingness to Interact

Our first result is that participants are significantly less likely to want to interact with counter-partisans than with co-partisans. Figure 3A shows that the average willingness to pay for interacting with one's partner is significantly lower in cross- than in co-partisan pairs. Moreover, Figure 3B shows that participants in cross-partisan pairs are significantly more likely to have a negative willingness-to-pay to interact, thus indicating a strict preference against interacting. Appendix Table A.6 shows that our results are robust to different specifications.[16] In a world of constrained time and attention, a greater willingness to interact with members of one's party implies the formation of echo chambers, personal networks or groups in which ideologically like-minded individuals are over-represented.

3.1.2. Barriers to Cross-Partisan Interactions

What motivates the observed gap in willingness to interact? In the context of our study, participants' willingness to pay for the conversation can be driven by both instrumental

[16] Appendix Table A.7 uses open-ended responses by participants to show that, across treatments, participants are not differentially apprehensive about technological or other disturbances, differentially suspicious of a conversation actually being implemented, or differentially struggling to understand the willingness to pay instructions. The table also suggests that these considerations do not loom large in the participants' minds to begin with.

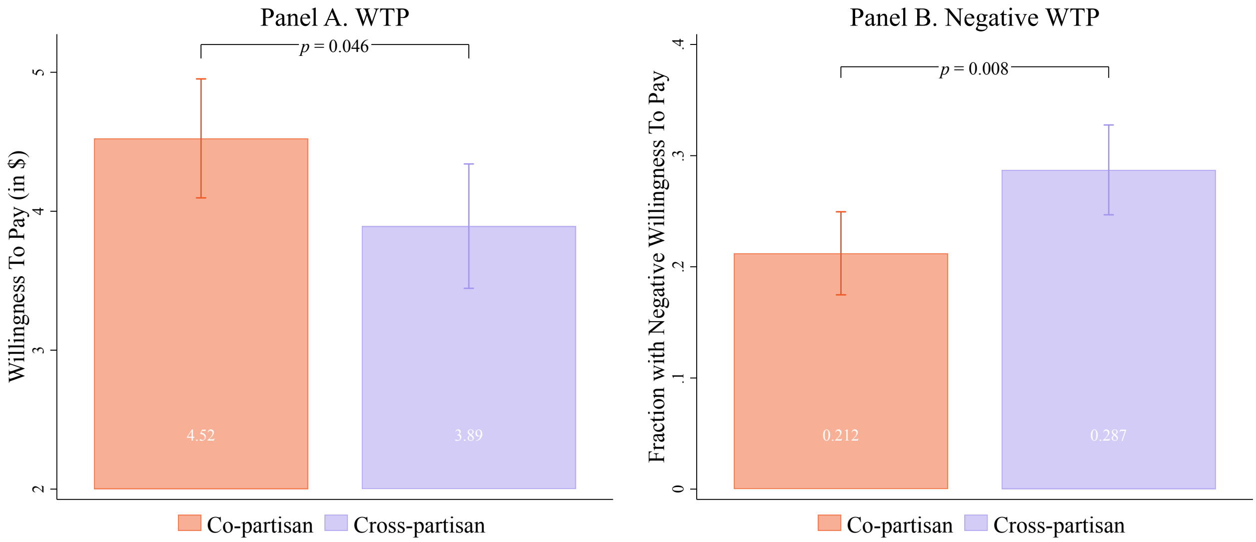

Figure 3: Willingness to Pay

Notes: The figure uses predicted values from regressions where Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using robust standard errors from relevant regressions.

and hedonic motives. Specifically, the conversation has instrumental value because it might help participants improve their scores on the quiz, and has hedonic value because it can be more or less pleasant, more or less interesting, etc.[17] The fact that $\sim$ 25 percent of participants have a strict preference against interacting is highly suggestive of the presence of hedonic motives (see Figure 3). Specifically, since information can always be ignored, its instrumental value is non-negative, implying that negative willingness-to-pay has to stem from hedonic factors.[18]

To better understand the drivers of participants' willingness to interact, we leverage data from an open-ended question asking participants for the rationale behind their choices in the willingness-to-pay elicitation. Participants' open-ended responses were hand-coded by a research assistant blind to treatment status. Figure 4A shows the four most frequently mentioned rationales and how they correlate with willingness to pay. The four most frequently mentioned reasons are the desire to improve on the quiz, curiosity, the expectation

[17]The conversation can also have instrumental value outside the context of our experiment, though such value is likely to be small. Nonetheless, our willingness-to-pay measure is designed to account for this possibility. Since virtually no participant mentions the instrumental value of the conversation outside the experiment when asked to explain the rationale behind her choices in the willingness-to-pay elicitation, we ignore such value in the remainder of the paper.

[18]Such hedonic factors, for instance, include unwillingness to talk to a stranger, unwillingness to discuss politics, etc.

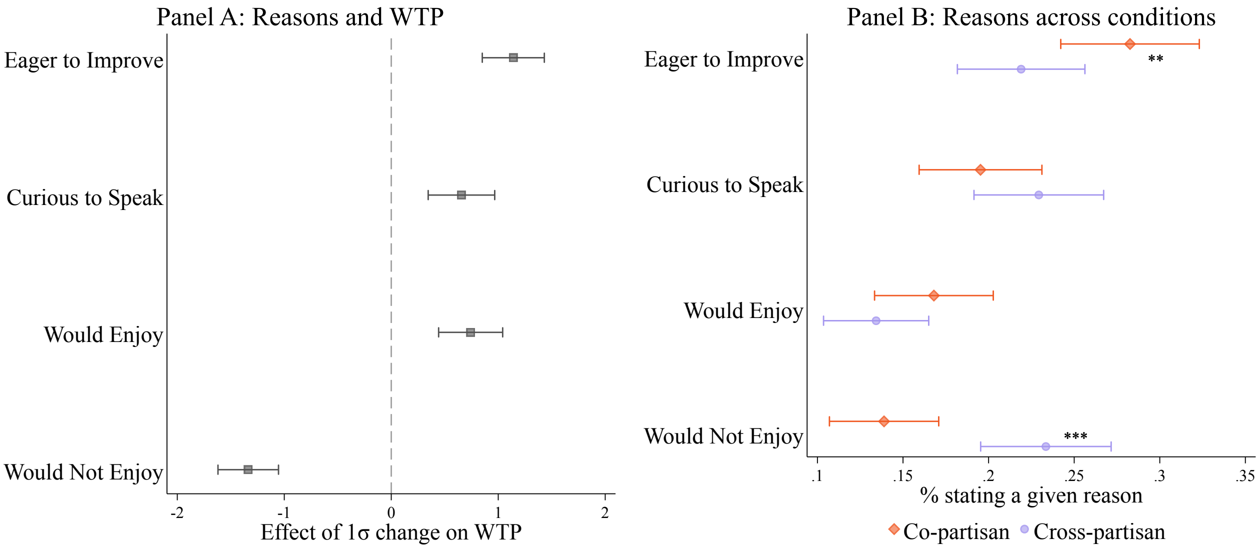

Figure 4: Reasons for WTP Decision

Notes: Panel A shows the coefficients from a regression of WTP (in US dollars) on the four most frequently mentioned rationales for the WTP decision. Panel B shows predicted values based on regressions of the four most frequently mentioned rationales for the WTP decision on a cross-partisan treatment indicator. Stars indicate the significant difference between co- and cross-partisan groups. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using robust standard errors from relevant regressions. * p<0.1, ** p<0.05, *** p<0.01.

that the conversation will be enjoyable, and the worry that the conversation will not be enjoyable.[19] These reasons correlate with willingness to pay for the conversation in intuitive ways. The desire to improve on the quiz, curiosity, and the expectation that the conversation will be enjoyable are all associated with a higher willingness to pay for the conversation. Conversely, worrying that the conversation will not be enjoyable is negatively correlated with willingness to pay.

Figure 4B shows how the four most frequently mentioned rationales behind the choices in the willingness to pay elicitation differ between experimental conditions. Participants poised to participate in cross-partisan interactions are significantly less likely to state that wanting to improve on the quiz factored into their willingness to pay for the conversation and significantly more likely to mention worries about the conversation being unpleasant. There are no statistically significant differences across conditions as to whether participants state that they are curious about the interaction or that they expect to enjoy it.[20]

[19]In Figure 4, we include separate indicators for participants mentioning expecting the conversation to be enjoyable and expecting it to be unenjoyable. This allows us to be consistent in the analysis presented in the figure by always constructing our outcome variables as indicators for mentioning a particular issue in the open response.

[20]Appendix Figure A.1 describes less frequently stated considerations about willingness to pay and illus-

In summary, participants' willingness-to-pay reflects both their desire to improve their quiz performance and concerns about not enjoying the conversation. These motives vary systematically between co- and cross-partisan interactions, aligning with the observed differences in willingness-to-pay.

In the next section, we drill down on participants' beliefs about the instrumental value of co- and cross-partisan conversations and compare these expectations to actual learning outcomes. In Section 3.3, we explore the hedonic aspects of the conversations.

3.2. The Instrumental Value of Conversations

A detailed investigation of the instrumental motives for selecting into echo chambers is important because, for consequential decisions, these motives might overshadow hedonic factors and be the primary drivers of self-selection into echo chambers.

3.2.1. The Gap in Expected Improvement

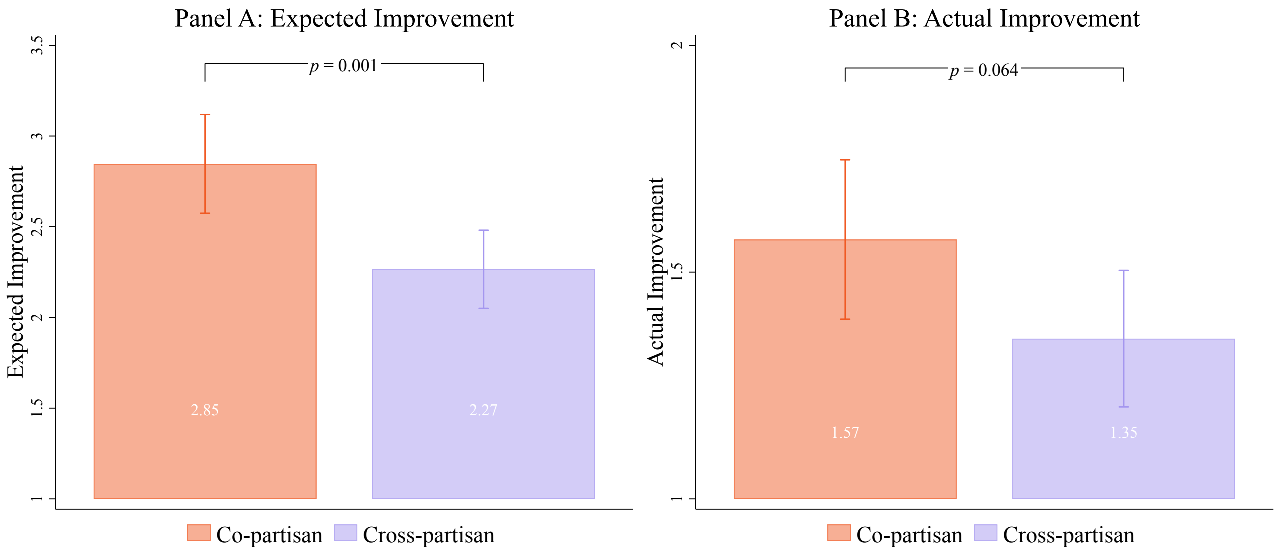

Figure 5: Expected and Actual Improvement

Notes: Panel A shows predicted values from a regression of expected improvement on a cross-partisan treatment indicator. Panel B shows the results from a regression of actual improvement on a cross-partisan treatment indicator. In the regressions underlying both panels, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using robust standard errors for Panel A, and standard errors clustered at the pair level for Panel B.

We begin our analysis of the instrumental value of co- and cross-partisan interactions by examining respondents' expectations about the extent to which the conversation can help

trates how such considerations vary by treatment.

them improve their answers to the quiz. If participants expect to learn less from counter-partisans than from co-partisans, then even absent hedonic factors such as the comfort of interacting with like-minded individuals, the perceived relative instrumental value of co- vs. cross-partisan conversations alone could lead to the formation of echo chambers.

As depicted in Figure 5A, participants expect to improve significantly less when they are assigned to a cross-partisan rather than a co-partisan conversation. The magnitude of the effect is around half a question in the quiz. As shown in Appendix Figure A.2, the conversation itself does not appear to alter these expectations significantly. When we assess participants' expected improvement after the conversation, we find that: a) they continue to expect greater instrumental benefits from co-partisan interactions compared to cross-partisan ones ($p = 0.015$), and b) the size of the expectation gap does not significantly change ($p = 0.250$). Thus, the conversation itself fails to significantly shift participants' beliefs about its instrumental value.

3.2.2. The Gap in Actual Learning and its Decomposition

To what extent are participants' expectations about learning from co- and cross-partisan conversations justified? In this section, we show that participants' relative pessimism about learning from cross-partisan interactions is qualitatively correct. We then provide a decomposition of participants' improvements on the quiz that helps us shed light on the reasons why participants in our experiment have a harder time learning from counter-partisans than from co-partisans.

We begin by analyzing the average differences in actual learning between co-partisan and cross-partisan conversations, where actual learning is defined as the increase in the number of correct answers to the quiz following the conversation. As shown in Panel B of Figure 5, we can be relatively confident that cross-partisan conversations are less informative than co-partisan ones. Specifically, we can reject at the 10 percent level the null hypothesis that co-partisan and cross-partisan conversations lead to the same level of improvement ($p = 0.064$). Even in the relatively unlikely event that cross-partisan conversations lead to larger improvements than co-partisan ones, we can rule out improvements greater than 0.01 correct answers on the quiz at the 5 percent significance level.[21] These results are qual-

[21] An improvement of 0.01 correct answers on the quiz translates to less than one cent in monetary terms.

itatively consistent with participants' expectations that co-partisan conversations are more informative than cross-partisan ones. Quantitatively the expected difference in improvements (0.58) is larger than the actual difference in improvements (0.22) between co- and cross-partisan conversations, albeit not significantly so at conventional levels ($p = 0.371$).

To understand why learning from co-partisans is by and large easier than learning from counter-partisans, we can decompose actual improvement into two key components: *potential for learning* and *difficulties in knowledge extraction*. We define potential for learning as the number of quiz questions for which a participant's partner knows the correct answer and the participant does not. We define difficulties in knowledge extraction as a participant's inability to correctly revise the answer to a question, conditional on her partner having the correct answer.

The decomposition of actual improvement into potential improvement and difficulties in knowledge extraction is helpful for shedding light on the reasons why participants are able to learn more from co-partisans than from counter-partisans. The potential for learning reflects differential knowledge distributions between co-partisan and cross-partisan pairs. Difficulties in knowledge extraction might stem, for instance, from a lack of trust in the credibility of the other side of the political aisle, a tendency of cross-partisan discussions to focus on unproductive questions, or a propensity of such conversations to become heated and hostile.

Formally, the decomposition works as follows. Let $improvement_{i,j}$ denote participant i's improvement on the quiz as a result of having a conversation with participant j. Let $potential_{i,j}$ denote the number of questions that participant i answered incorrectly in the initial quiz and participant j answered correctly. By the law of total expectation, we have

$$E(improvement_{i,j}) = \sum_{k=0}^{14} E(improvement_{i,j}|potential_{i,j} = k)P(potential_{i,j} = k)$$

where the expectations indicate population-level quantities.

Thus, even at the upper bound of the confidence interval for the relative expected improvement from cross-partisan conversations, the monetary gain is so minimal that it would not alter the behavior of any participant in our experiment who, according to her willingness-to-pay, has a strict preference against interacting with a counter-partisan. In other words, even if participants' beliefs about the relative gains from cross- versus co-partisan conversations were mechanically set at the upper limit of the confidence interval, this marginal improvement would not be sufficient to shift their preference away from self-selecting into echo chambers.

In order to produce one summary measure of perceived difficulties in knowledge extraction, we impose the assumption that, conditional on player j correctly answering k questions in the initial quiz that player i answered incorrectly, player i improves her score on the Revised Quiz by, on average, βk questions. We can thus interpret $\beta \in [0,1]$ as parametrizing the perceived ease of knowledge extraction, and we can let it vary by treatment $t \in \{co, cross\}$. The assumption allows us to rewrite expected improvement as:

$$E(improvement_{i,j}) = \beta_t E(potential_{i,j}) \tag{3.1}$$

By conditioning the expectations in the expression above on whether participant i is poised to have a conversation with a co- or counter-partisan, we obtain four measures: two measures of potential improvement (one from co- and one from counter-partisans), and two measures of ease of knowledge extraction (again, one from co- and one from counter-partisans).

Panels A and B of Figure 6 show that impairments to learning in cross-partisan interactions are primarily due to difficulties in knowledge extraction. Specifically, Panel A of Figure 6 shows that potential improvement is, if anything, larger in cross-partisan conversations than in co-partisan ones. This is because, as shown in the next section, knowledge is distributed across party lines, with Democrats and Republicans being equally informed on average, but differentially informed across quiz questions. The fact that, despite the greater potential for learning in cross-partisan conversations, participants tend to learn less in cross-partisan than in co-partisan interactions must then be due to greater difficulties in knowledge extraction when meeting counter-partisans. Indeed, Panel B of Figure 6 shows that, for every question that a participant's conversation partner answers correctly and the participant does not, the participant's score improves by 17 percent less when the partner is a counter-partisan rather than a co-partisan. We explore the reasons behind the difficulties in knowledge extraction in Section 3.2.4.

Our definition of potential improvement and our decomposition of learning into potential improvement and knowledge extraction make two implicit assumptions. The first assumption is that the most valuable learning opportunities are instances where exactly one of the two conversation partners has the correct answer, as opposed to instances where neither partner has the correct answer. The second assumption is that unlearning, defined

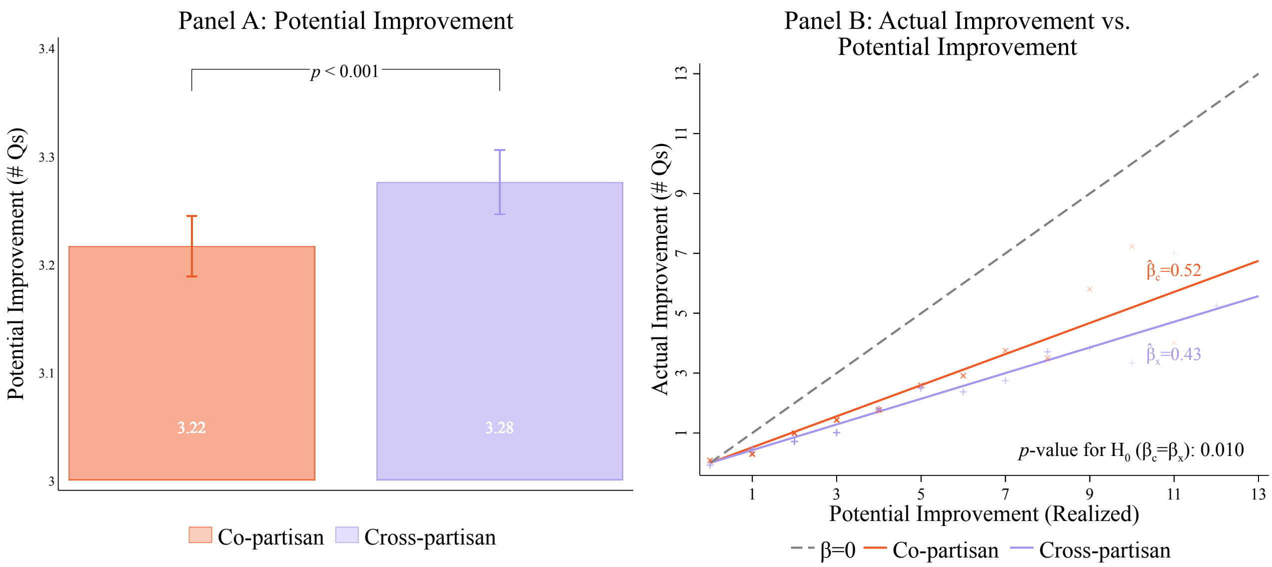

Figure 6: Actual and Potential Improvement

Notes: Panel A reports potential improvement by treatment. Potential improvement is constructed as follows. First, we select two parties i and j from the set $\{Democrat, Republican\}$. Second, for each question, we calculate the potential improvement for the average affiliate of party i who meets with the average affiliate of party j by taking the sample-level probability that the affiliate of party i does not know the answer to the question times the probability that the affiliate of party j does. These question-level statistics are then summed up at the quiz level and averaged across cross-partisan and co-partisan interactions to produce the bars in the graph. In Panel B, we relate potential improvement at the level of individual pairs to actual improvement in those pairs. We refer to our measure of potential improvement in Panel B as potential improvement (realized), because it refers to the number of questions, for each participant and for the partner she is randomly matched to in conversation, where the participant gave an incorrect answer and her partner gave the correct one. Panel B presents the results of regressing actual improvement for each participant on potential improvement (realized), in a model where the intercept is fixed at 0 as required by Equation 3.1. In the regression, potential improvement (realized) is interacted with treatment assignment. β_c and β_x represent the estimated ease of knowledge extraction for the Co- and Cross- partisan treatments respectively. In the overlayed scatterplot of panel B, Republican/Democrat-only pairs are weighted when calculating the average of actual improvement within each value of potential improvement to address sample imbalances; the opacity is weighed by the square root of the number of observations in each bin. We employed the square root of the number of observations in each bin rather than the raw number of observations, because some of the coefficients would have been hardly visible otherwise. In all regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. All outcomes are expressed in terms of number of questions on the quiz (out of 14). The 95 percent confidence intervals and p-value in Panel A are computed using bootstrapped standard errors; the ones in Panel B are computed using standard errors clustered at the pair level.

as transitioning from having a correct answer to having an incorrect one as a result of the conversation, is sufficiently uncommon as to be negligible.

Appendix Table A.8 provides support for both assumptions. First, instances in which conversation partners transition from one correct answer between them to two correct answers are 4 to 5 times more common than instances in which the partners transition from zero correct answers between them to one correct answer.[22] Moreover, transitioning from one correct answer between the two partners to two correct answers is 16 times more likely than transitioning from zero correct answers to two correct answers. Second, we do not find evidence that unlearning plays a significant role in our experiment. Specifically, the fraction of instances in which at least one partner transitions from a correct answer to an incorrect

[22]Here an "instance" is a particular question in which a particular participant pair transitions from one knowledge state to another.

one is less than 6 percent.[23]

Mirroring our results on treatment differences in knowledge extraction, Appendix Table A.8 also shows treatment differences in exactly two of the nine possible pairwise transitions: cross-partisan conversations are significantly less likely to feature transitions from one correct to two correct answers and they are significantly more likely to lead to stagnant transitions from one correct answer to one correct answer. We explore this further in Section 3.2.4.

3.2.3. My-side Bias and Distributed Knowledge

We showed that participants hold pessimistic beliefs about the extent to which they can learn from counter-partisans and that such beliefs are qualitatively correct. In this section, we show that, despite participants' beliefs about average learning being largely accurate, they are likely to be so, at least in part, for the wrong reasons. Specifically, participants hold unduly pessimistic beliefs about both the average knowledge of counter-partisans and the degree of complementarity between what counter-partisans know and what they know.

For every question on the quiz, we elicited participants' beliefs about the proportion of Republicans and Democrats who, in a pilot version of the study, answered the question correctly. These beliefs are depicted in panels A and B of Figure 7. As shown in the figure, both Republicans and Democrats exhibit substantial my-side bias:[24] they believe that, on average, co-partisans are more informed than counter-partisans, with a 6.44 percentage point overestimation of how many more correct answers co-partisans had in the initial quiz.[25] In contrast to participants' beliefs, we find no gap in the actual average informedness across the two parties.

[23] These results suggest that people who give the correct answer in the Initial Quiz are not simply guessing randomly. In fact, the results presented in Panel B of Figure 6 are robust to different specifications of potential learning in which one's conversation partner is considered to know the correct answer only if she also reports a high level of confidence in her answer (see Appendix Figure A.3).

[24] Here we use 'bias' to mean 'partiality toward' rather than 'behavioral bias.' In fact, in a world of heterogeneous priors, it is consistent with rationality for individuals who choose a party based on their prior to expect their party's supporters to be better informed than those of the opposing party.

[25] This number is obtained by taking the average degree of overestimation by Democrats across all 14 questions, summing it to the average degree of overestimation by Republicans across all 14 questions, and dividing the result by two.

Participants not only overestimate the average knowledge of their co-partisans relative to counter-partisans, but also overlook the fact that knowledge is distributed across party lines, meaning that Democrats are more likely to know the answers to some of the questions that Republicans struggle with, and vice versa. Specifically, both Democrats and Republicans believe the there is no single question on the quiz where, on average, counter-partisans perform significantly better than co-partisans. In contrast to participants' beliefs, we find that: i) Democrats are significantly more likely than Republicans to correctly answer the questions about inflation, the identity of the Speaker of the House, the deportation of immigrants, and Joe Biden's stance on the Defund the Police movement; ii) Republicans are significantly more likely than Democrats to answer correctly the three questions about the prevalence of background checks for gun sales, Donald Trump's official stance on vaccines, and the number of illegal immigrants in the United States.[26]

My-side bias and disbelief in distributed knowledge lead to a lower perceived potential for learning from cross-partisan conversations. This perception sharply contrasts with the greater actual potential for learning from counter-partisans, as documented in the previous section. Appendix Figure A.4 illustrates the significant gap between perceived and actual learning potential.

In sum, although participants accurately assess overall learning, their belief that co-partisan conversations are more informative than cross-partisan ones stems, in part, from a mistaken view of the potential to learn from the other side of the political aisle. This misconception remains largely unaffected by the experience of cross-partisan interactions, which are no more effective than co-partisan ones in correcting misperceptions about the distribution of knowledge (see Appendix Figure A.5).

3.2.4. What Drives the Greater Difficulty of Extracting Knowledge from Counter-Partisans?

Figure 6 presented evidence that participants face greater difficulties in knowledge extraction when they meet counter-partisans than when they meet co-partisans. This section explores why knowledge extraction might be harder in cross- than in co-partisan conversa-

[26] The experiment took place in early 2024, when Donald Trump had a less ambiguous stance regarding vaccines than he did later on.

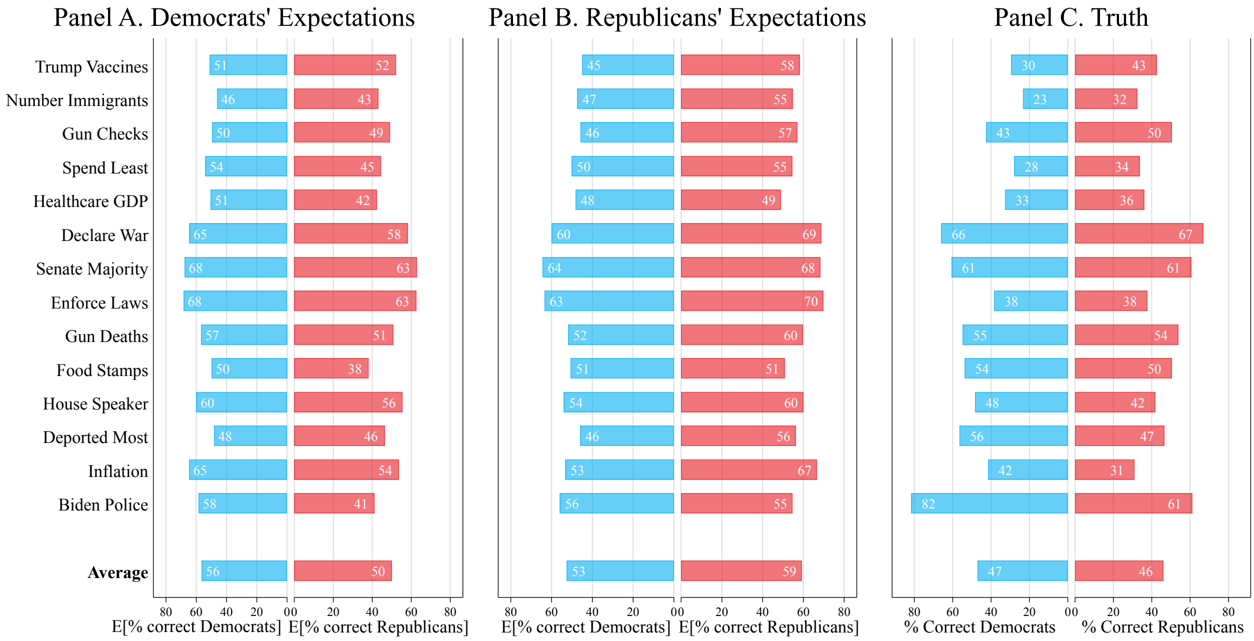

Figure 7: Expected and Actual Shares of Correct Responses

Notes: The figure presents the expected and actual shares of correct responses by respondents affiliated with different parties. Each row refers to a question from the factual quiz (the crosswalk from labels in this figure to full questions can be found in Appendix Table A.1). The rows are ordered by first calculating the difference between the share of correct answers among Republicans and the share of correct answers among Democrats in the Initial Quiz, and then by sorting questions in decreasing order by that difference. Panel A displays the expectations of Democrats about other Democrats and about Republicans. Panel B displays the expectations of Republicans about Democrats and about other Republicans. Panel C reports the actual shares of correct responses to the Initial Quiz of Republicans and Democrats in our main sample.

tions. Our empirical strategy is laid out in Appendix C and proceeds as follows. First, we ask whether a given feature of conversations or of its participants varies by treatment, in the sense of differing between co- and cross-partisan conversations. Second, we ask whether the feature in question predicts knowledge extraction. Specifically, we ask whether the feature moderates the rate at which potential improvement translates into actual learning, the measure of knowledge extraction we defined in Section 3.2.2. Third, we deem a feature of the conversation a possible driver of the gap in knowledge extraction if it is higher (lower) in cross-partisan conversations and predicts lower (higher) knowledge extraction.

The first mechanism we explore relates to participants' perceptions of the credibility of co- and counter-partisans as information sources. In the previous section, we showed that participants generally expect counter-partisans to be less informed than co-partisans, and that they underestimate the extent to which knowledge is distributed across party lines. In addition to these pessimistic initial beliefs, encountering disagreement during cross-partisan interactions may further erode trust in counter-partisans as credible sources of information. This is in line with a mechanism similar to the one proposed by Gentzkow and Shapiro (2006), where people assess the credibility of an information source based on how closely the signal emitted by the source aligns with their own prior beliefs. Since participants are more likely to face disagreement in cross-partisan conversations than in co-partisan ones, it is unlikely that cross-partisan interactions will dispel participants' pessimistic expectations.[27,28]

In line with the reasoning above, we find that the conversations themselves do not significantly improve participants' beliefs about the credibility of counter-partisans as information sources. Specifically, Appendix Table C.1 shows that, after the conversation: i) participants in cross-partisan interactions believe that their partners answered fewer questions correctly than participants in co-partisan interactions, and ii) participants in cross-partisan conversations are less likely to agree with the statement that their partner 'knew more than they did.' Crucially, the table also shows that higher levels on either measure of trust in a partner's knowledge are associated with greater knowledge extraction.

Appendix Figure A.6 provides a behavioral counterpart to the beliefs above. Specifically, it shows that, conditional on potential improvement, participants are more likely to modify their answers after conversations with co-partisans than with counter-partisans. In other words, when participants observe a counter-partisan give a different answer to a question, they are less inclined to modify their own response than they would be if the different answer came from co-partisans. The pattern of results from the last three paragraphs suggests that perceived credibility plays a significant role in actual learning outcomes.

Appendix C also explores psychological factors that might underlie difficulties in knowledge extraction from counter-partisans. To this end, we analyzed the video recording of the conversations and we measured participants' subjective experience of and emotions pertaining to the interaction. We find that the psychological factors we initially hypothesized would make knowledge extraction harder in cross-partisan conversations did not seem to play a major role. For example, we expected cross-partisan conversations to be more likely

[27]It is worth noting that the initial pessimistic beliefs about counter-partisans' credibility may themselves stem from a version of the mechanism from Gentzkow and Shapiro (2006) operating outside the context of our experiment.

[28]Appendix Table B.1 provides the interesting descriptive finding that participants spend more time discussing the answers on which they disagree, and it shows that this pattern is especially pronounced in cross-partisan conversation. This evidence rules out the possibility that learning gaps across treatments result simply from avoiding discussing issues where there is disagreement.

to become heated and confrontational. However, our analysis of the video interactions shows that heated conversations are rare, that they are no more common in cross-partisan than in co-partisan interactions, and that signs of animosity do not significantly predict knowledge extraction. Furthermore, self-reports describing the conversation as ’confrontational’ or ’a good time,’ as well as emotions like aversion, anxiety, or enthusiasm, do not explain differences in knowledge extraction across treatments. Finally, we found no evidence that factors that research in psychology suggests aid persuasion (e.g., speech speed, filler word usage, and a declining intonation at the end of sentences) predict knowledge extraction in our setting (Smith and Clark, 1993; Guyer et al., 2019). In Appendix B, we describe how several of the variables above were extracted from the video data and provide some descriptives of the conversations’ structure and focus.

3.2.5. Instrumental Value of the Conversation: Taking Stock

In this section, we studied the perceived and actual instrumental value of selecting in and out of echo chambers. The key insights from this analysis are i) that participants perceive the instrumental value of co-partisan conversations to be higher than that of cross-partisan ones, ii) that although some components of these expectations are qualitatively mistaken (e.g., participants exhibit my-side bias), participants’ overall beliefs that co-partisan conversations are more informative than cross-partisan ones are qualitatively correct, iii) that the main reason behind the discrepancy in learning from co- and cross-partisan conversations is that knowledge extraction is significantly more difficult in cross-partisan interactions, and iv) that such differences in knowledge extraction seem to primarily be due to differences in the perceived credibility of co- and counter-partisans as information sources.

An additional result about how the conversations change participants’ answers provides further context to our findings above. In Appendix F, we show that, relative to co-partisan conversations, cross-partisan conversations lead to less stereotypical beliefs. Specifically, before the conversation, a participant’s profile of answers carries a clear partisan signature. In particular, we are able to correctly predict someone’s party affiliation by looking at their answers 63 percent of the time. After engaging in a cross-partisan conversation, this fraction drops to 47 percent. Moreover, this factual depolarization persists in the follow-up survey.

Our observation that cross-partisan contact makes factual beliefs less stereotypical of a participant's party has some analog in previous work on cross-partisan contact that finds that contact leads to a convergence of attitudes and beliefs (e.g. Hobolt et al. 2024). However, our design is unique in allowing us to observe participants' factual beliefs vis-à-vis an unambiguous ground truth. This, in turn, allows us to study convergence to the truth as a phenomenon that may be distinct from factual depolarization. Leveraging this distinction, we see that cross-partisan conversations might be worse at achieving learning than at achieving factual depolarization.

3.3. The Hedonic Value of Conversations

3.3.1. Expected and Realized Hedonic Value

In Section 3.1.2, we showed that participants were significantly more apprehensive when anticipating a conversation with a counter-partisan than one with a co-partisan. Specifically, facing cross-partisan conversations, they were more likely to voice the concern that the conversation would not be enjoyable. Panel A of Figure 8 confirms this result in a test that subsumes positive and negative mentions of predicted enjoyment (as well as the absence of any mentions) in an index.

Panel B of Figure 8 reveals that participants' worries about the relative unpleasantness of cross-partisan contact are largely unwarranted. The figure depicts agreement with the statement "I had a good time during the conversation" on a Likert scale from 1 (strongly disagree) to 4 (strongly agree). We see that the average participant in both conversations experienced conversations as rather pleasant and that participants were equally likely to report having had a good time in co-partisan and cross-partisan conversations.

Unlike predictions about expected learning, predictions about enjoying the conversation are not incentivized; furthermore, they are measured on a scale that is different from the one participants used ex-post to report the extent to which they enjoyed the conversation. Thus, we have to interpret these results with caution. Nonetheless, we argue that participants are overly pessimistic about the *relative* hedonic value of co- and cross-partisan conversations, because the ex-ante and ex-post differences in enjoyment exhibit qualitatively different patterns. Specifically, the ex-ante difference in the likelihood of voicing concerns

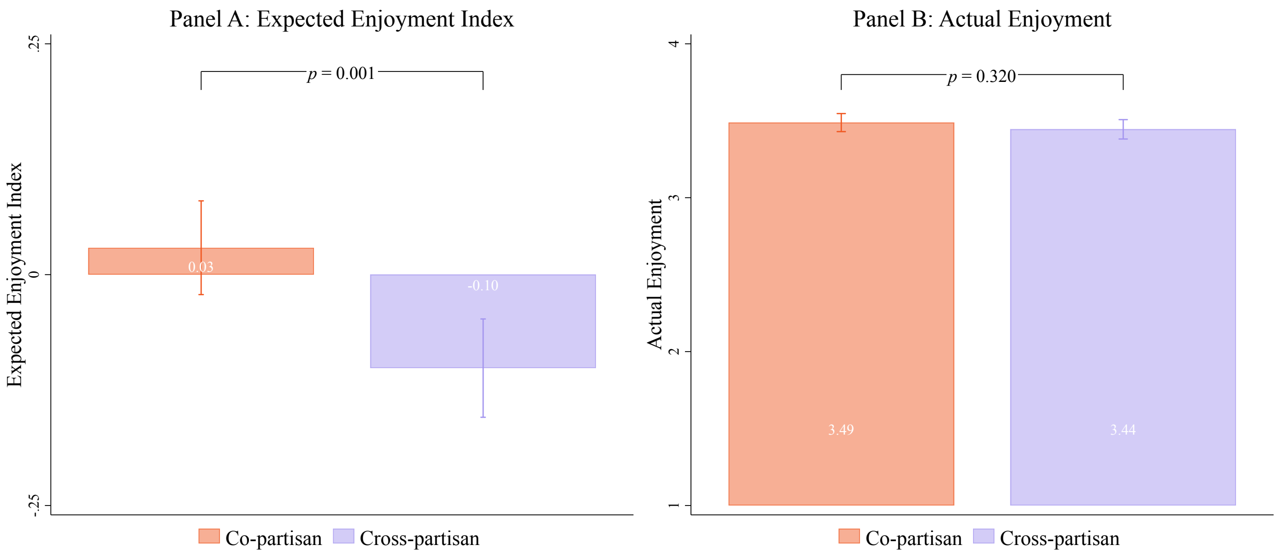

Figure 8: Expected and Actual Enjoyment

Notes: Panel A reports the predicted values of a regression of an index of expected enjoyment on an indicator for being in the cross-partisan treatment. The expected enjoyment index is coded on a scale from -1 to 1, with -1 indicating that the participant reported expecting not to enjoy the interaction, 1 indicating that the participant reported expecting to enjoy the interaction, and zero indicating that the participant did not mention anything related to enjoyment in the open-ended question asking about the reasons behind her willingness-to-pay decision. Panel B reports the predicted values of regressing reported enjoyment after the interaction on an indicator for being in the cross-partisan treatment. The outcome in Panel B uses a Likert scale from 1 to 4, with 1 being "strongly disagree" and 4 being "strongly agree." In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals and p-values in the figure are computed using robust standard errors in Panel A and using standard errors clustered at the pair level in Panel B.

about not enjoying the conversation is significantly negative, whereas we see no ex-post difference in self-reported experiences.

These suggestive results dovetail with the significant effects of conversations on affective polarization that we describe next.

3.3.2. Affective Polarization: Short-term Effects

We now examine whether cross-partisan conversations improve how participants feel about individuals affiliated with the opposing political party — i.e., whether these conversations reduce affective polarization. Given the recent rise in affective polarization and its well-documented negative effects on democratic processes and social cohesion, any reduction in affective polarization carries significant normative weight (Iyengar et al., 2019; Boxell et al., 2024).

Consistent with prior findings on the pronounced levels of affective polarization in the U.S. political landscape (e.g. Druckman and Levy, 2022), our primary measure of affective polarization, the feeling thermometer, illustrates substantial polarization at baseline. Par-

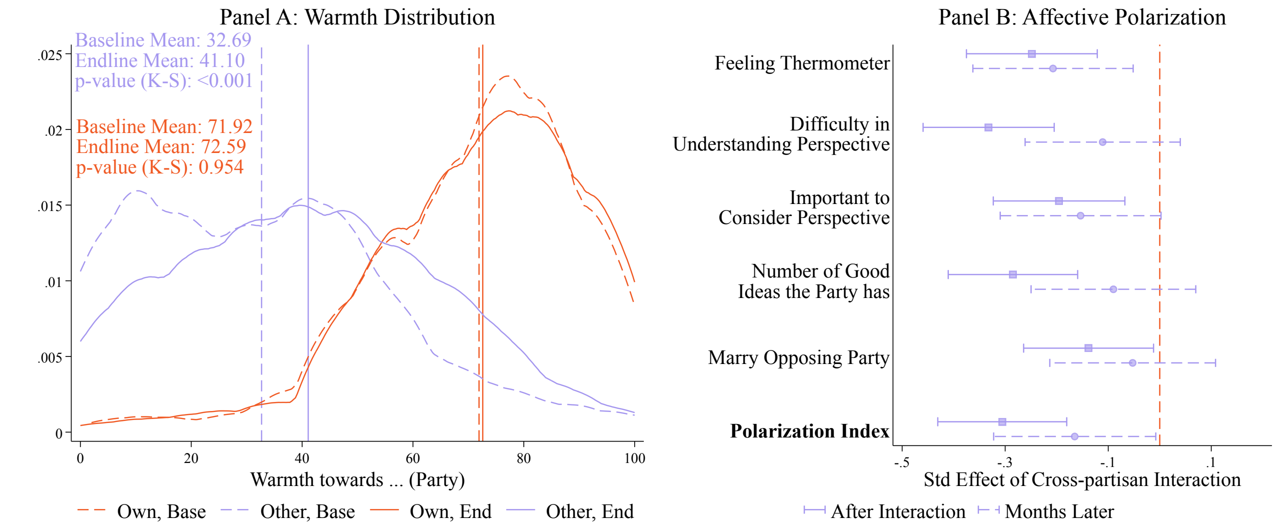

Figure 9: Affective Polarization

Notes: Panel A shows the distributions of baseline and endline warmth towards one's party and the opposing party, restricting the sample to the cross-partisan treatment group. The p-value from a Kolmogorov-Smirnov test for equality of distribution of the warmth towards one's own party before and after the interaction is 0.954. The p-value from a Kolmogorov-Smirnov test for equality of distribution of the warmth towards the other party before and after the interaction is <0.001. Panel B plots coefficients on an indicator for being assigned to a cross-partisan interaction from regressions of the 5 outcome variables in the figure — together with an index of affective polarization constructed as described in Section 2.2 — on the indicator. These outcomes are measured twice, after the interaction and, again, approximately 100 days after the interaction. We deviate from the pre-registration by including the "Marry opposing party" variable in the polarization index. This provides a more comprehensive and more conservative picture of our results as discussed in the main text. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using standard errors from relevant regressions clustered at the pair level.

ticipants rated their feelings toward individuals from the opposing party 39.2 points colder (on a scale from 0 to 100) compared to those from their own party.

Does our intervention mitigate this polarization? The evidence suggests it does. First, Figure 9A shows, in a before-after comparison, that cross-partisan conversations reduce the feeling thermometer gap by $\sim$ 20 percent, indicating a notable improvement in attitudes toward the opposing party. Second, we employ five distinct measures of affective polarization elicited after the conversations to estimate the differential effects of cross-partisan versus co-partisan interactions. Panel B of Figure 9 demonstrates that cross-partisan conversations produce significant and substantial reductions in affective polarization across all five measures, with effect sizes ranging from 0.14 to 0.33 standard deviation units.

The findings above suggest that cross-partisan conversations not only reduce explicit coldness toward the opposing party but also positively shift other polarization-related attitudes. This is suggestive of the potential for cross-partisan fact-based conversations about politics to serve as a tool for bridging partisan divides.

3.3.3. Persistence

Although our intervention shows promise as a tool to promote social cohesion, two important concerns warrant further attention. First, the transparency of the experiment may raise the possibility of experimenter demand effects, wherein participants modify their behavior to align with their perceptions of the experimental hypotheses. Second, one might worry that the observed effects may be short-lived, diminishing their practical significance.

To address these concerns, we conducted a follow-up survey approximately 100 days after the intervention.[29] As discussed in Section 1, the survey was obfuscated; i.e., it was designed to appear unrelated to the original experiment. This approach minimizes the likelihood that participants associate the follow-up survey with the initial study, thereby substantially reducing the risk of experimenter demand effects. We note that the follow-up survey does not allow us to fully disentangle the impact of mitigating demand effects from the persistence of treatment effects over time. Our design, however, effectively bounds experimenter demand effects in the short-term.[30]

The primary outcome measure specified in our pre-analysis plan, the feeling thermometer, indicates that the treatment effect of cross-partisan interactions remains positive and statistically significant in the follow-up survey ($p = 0.009$). This long-term effect is 92 percent of the immediate effect, and the difference between the immediate and follow-up effects is not statistically significant at conventional levels ($p = 0.541$). Figure 9B shows that the persistence of the effects varies across different outcomes, but that the overall pattern remains consistent. Specifically, on a pre-specified equally weighted index of our five standardized affective polarization measures, we observe a 0.15 standard deviation units reduction in polarization in the follow-up survey ($p = 0.058$) when comparing participants in cross-partisan pairs to those in co-partisan pairs.[31]

Taken together, our results suggest that the effects of cross-partisan conversations on

[29] Appendix Table A.2 shows no differential response rates across treatments.

[30] Under the plausible assumption that the effects of the intervention do not increase over time.

[31] In the pre-analysis plan we indicated that this index would exclude the fifth polarization outcome, already found to be less malleable in Levy et al. (2022). If we followed the pre-analysis plan and excluded this outcome, the gap at follow-up would be larger (0.17 standard deviations) and significant at the 5 percent level ($p = 0.040$). On a separate note, Appendix Figure A.7 shows qualitatively similar results when restricting the analysis to participants observed both in the main study and the follow-up survey.

reducing affective polarization are both substantial and largely persistent. Even several months after the intervention, participants continue to exhibit reduced levels of polarization, indicating that the impact of engaging in factual political discussions extends well beyond the immediate context of the experiment.

3.3.4. Relation to the Literature on Interpersonal Contact

We contextualize our affective polarization results by comparing them to those in the literature on interpersonal contact, which has traditionally emphasized interactions between different racial, ethnic, or social groups.[32] In contrast to existing review articles, our analysis: i) narrows its focus to experimental studies that involve actual interpersonal contact (as opposed to hypothetical or educational interventions), ii) includes both studies conducted in-person and online, and iii) extends the scope to include contact across ideological divides.

The standardized treatment effects we report — especially after correcting for potential publication bias — are broadly comparable to the meta-analytic effects of interpersonal contact documented in the literature (see Appendix G for details on our meta-analysis methodology). We note that our study is among the largest in this literature and among the few that involve an obfuscated follow-up study to measure the persistence of treatment effects (see Appendix Figure G.1). It is also one of the only studies to measure expectations about the contact experience, to identify prejudice by contrasting those expectations with the realized experience, and to study selection into the contact with out-group members. We discuss this last point in the next section.

3.3.5. Selection Effects on Affective Polarization Reduction

A difficulty with assessing the external validity of various studies in the literature on inter-group contact, including several featured in the meta-analysis discussed above, is that individuals who particularly dislike contact with out-group members can select out of the study. This raises the question of whether individuals who select into the study are more receptive to the depolarizing effects of inter-group contact than those who choose to opt-out. In our study, not only can we show that there is no effect of being assigned to a cross-partisan con-

[32]See Paluck et al. (2021) for an older but comprehensive survey of this literature and Lowe (2024) for a selective survey of studies that had been pre-registered on the AEA-RCT Registry and the EGAP Registry.

versation on selection out of the study, but we can also explore how the depolarizing effects of cross-partisan interactions depend on participants' eagerness to engage in them.

In order to study whether participants who are more eager to engage in cross-partisan contact experience greater reductions in affective polarization, we relate a participant's willingness to pay for a cross-partisan conversation to the individual-level before-after reduction in affective polarization estimated using the feeling thermometer. We find that the correlation between the two measures is not statistically different from zero ($corr = 0.013$, $p = 0.772$). We interpret this result as assuaging concerns that the positive effects of cross-partisan contact on affective polarization are more pronounced for individuals who are more willing to engage with counter-partisans.

3.3.6. Predictors of Affective Depolarization in Cross-partisan Conversations

Lastly, we turn to the question of which features of a cross-partisan conversation or of its participants are predictive of affective depolarization. In Appendix D, we describe the results of a set of analyses that regress individual-level before-after differences in affective polarization on various observables.

We find that the depolarizing effect of cross-partisan contact is greater among female and younger participants, as well as those who are more polarized at baseline or score lower on a social potency scale.[33] Notably, the greater depolarization among participants with higher baseline affective polarization indicates that our intervention is not merely "preaching to the choir" but can also reach and influence more polarized individuals. Meanwhile, the association between higher social potency and reduced depolarization suggests that personality traits linked to resistance to persuasion may hinder the positive effects of cross-partisan contact.

Other characteristics of the conversation and of one's conversation partner also play a role in shaping the depolarizing effects of cross-partisan interactions. For instance, being paired with an older conversation partner is associated with greater depolarization. In contrast, and perhaps surprisingly, neither the partner's actual knowledge nor the participant's perception of their partner's knowledge predicts depolarization. The presence of animosity or confrontational behavior during the conversation is also not correlated with depolariza-

[33]Social potency refers to a tendency to strive for interpersonal dominance.

tion, possibly because heated and confrontational interactions are rare in our context. By contrast, participants who reported having a "good time" and experienced more enthusiasm and less aversion during the conversation exhibited greater depolarization.

Although this evidence is correlational rather than causal, it can be helpful by shining a light on potential drivers of the reduction in affective polarization.

3.3.7. Hedonic Value of the Conversation: Taking Stock

In this section, we explored the hedonic motives that influence participants' engagement in co-partisan and cross-partisan conversations. Participants expected conversations with co-partisans to be more enjoyable than those with counter-partisans, but we found no evidence that, ex-post, co-partisan interactions were rated as more enjoyable than cross-partisan ones.

The hedonic impact of conversations goes beyond mere enjoyment. One of the most striking results of our analysis is the significant reduction in affective polarization following cross-partisan conversations. Immediately after the experiment, participants showed a notable improvement in their attitudes toward those from the opposing party, with a reduction of approximately 20 percent in the feeling thermometer gap. Moreover, this reduction in affective polarization persisted: approximately 100 days after the intervention, the treatment effects remained substantial, retaining 92 percent of the immediate reduction. Notably, the reduction in affective polarization was not limited to participants who were more eager to engage in cross-partisan contact, and is associated with a host of intuitive factors such as the degree to which the conversation partners reported enjoying the interaction. These results highlight the value of engaging in structured, factual conversations across party lines as a promising tool for reducing polarization and improving social cohesion.

4. Conclusion

Our experiment delivers new evidence on the drivers and consequences of co- and cross-partisan conversations centered on facts about politics. We identify a preference for co-partisan over cross-partisan conversations that might contribute to the prevalence of echo chambers even where geographical sorting offline and algorithmic sorting online are not at play. Our finding that both hedonic and instrumental motives play a role in driving the preference for co-partisan interactions likely makes self-selection into such interactions more robust to the exact nature and objective of the interaction than it would otherwise be.

A nuanced picture emerges when we look at the consequences of cross-partisan contact. In terms of information aggregation, we document significant difficulties in extracting knowledge from counter-partisans: our participants struggle to harness the benefits of knowledge that is distributed across the aisle. As a result, policies that encourage cross-partisan interactions may not significantly enhance cross-partisan information sharing. Because the conversations do little to change broadly correct beliefs about overall learning and mistaken beliefs about the informedness of counter-partisans, policies that induce cross-partisan interactions are also not likely to reduce future self-selection on informational grounds. At the same time, our results suggest the intriguing possibility that policies that correct prejudice about the other side's informedness can both increase the willingness to engage in cross-partisan contact and improve how much individuals learn from cross-partisan contact. Future work could put the potential of such policies to foster cross-partisan learning to an experimental test.

In terms of social cohesion, our findings suggest that encouraging cross-partisan interactions may lastingly promote social cohesion and reduce partisan divides. This stands in contrast to evidence in Santoro and Broockman (2022) showing that conversations about political ideology more broadly do not have a depolarizing effect. Our results suggest that it is neither the absence of disagreement nor the focus on something other than politics that are necessary conditions for cross-partisan contact to be depolarizing. Instead, depolarization may be aided by a focus on facts and by participants' incentives being broadly aligned (Lowe, 2021). In particular, these facets of our interaction environment may help participants avoid "digging in their heels" and self-persuading in an effort to persuade their counterpart (Schwardmann and Van der Weele, 2019; Schwardmann et al., 2022). Future work could vary the topics and incentives associated with conversations to shed further light on the exact conditions under which cross-partisan contact depolarizes.

References

Acemoglu, Daron, Cevat Giray Aksoy, Ceren Baysan, Carlos Molina, and Gamze Zeki, "Misperceptions and Demand for Democracy under Authoritarianism," Technical Report, National Bureau of Economic Research 2024.

Allport, Gordon Willard, *The Nature of Prejudice*, Addison-Wesley, 1954.

Andrews, Isaiah and Maximilian Kasy, "Identification of and Correction for Publication Bias," *American Economic Review*, August 2019, *109* (8), 2766–2794.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson, "The Diffusion of Microfinance," *Science*, 2013, *341* (6144), 1236498.

_ , Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials," *The Review of Economic Studies*, 02 2019, *86* (6), 2453–2490.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya, "Facts, alternative facts, and fact checking in times of post-truth politics," *Journal of public economics*, 2020, *182*, 104123.

Bauer, Kevin, Yan Chen, Florian Hett, and Michael Kosfeld, "Group identity and belief formation: a decomposition of political polarization," 2023.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong, "Unity in Diversity? How Intergroup Contact Can Foster Nation Building," *American Economic Review*, November 2019, *109* (11), 3978–4025.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Belot, Michèle and Guglielmo Briscese, *Bridging America's Divide on Abortion, Guns and Immigration: An Experimental Study*, Centre for Economic Policy Research, 2022.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 1992, *100* (5), 992–1026.

Bishop, B. and R.G. Cushing, *The Big Sort: Why the Clustering of Like-minded America is Tearing Us Apart*, Houghton Mifflin, 2008.

Blattner, Adrian and Martin Koenen, "Does Contact Reduce Affective Polarization? Field Evidence from Germany," SSRN Electronic Journal July 2023.

Boisjoly, Johanne, Greg J. Duncan, Michael Kremer, Dan M. Levy, and Jacque Eccles, "Empathy or Antipathy? The Impact of Diversity," *American Economic Review*, December 2006, *96* (5), 1890–1905.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro, "Cross-country trends in affective polarization," *Review of Economics and Statistics*, 2024, *106* (2), 557–565.

Braghieri, Luca, "Political Correctness, Social Image, and Information Transmission," *American Economic Review. Forthcoming*, 2024.

_ , Sarah Eichmeyer, Ro'ee Levy, Markus Möbius, Jacob Steinhardt, and Ruiqi Zhong, "Article-Level Slant and Polarization of News Consumption on Social Media," *Working Paper*, 2024.

Brown, Jacob, Enrico Cantoni, Ryan Enos, Vincent Pons, and Emilie Sartre, "The Increase in Partisan Segregation in the United States," *Working Paper*, 2024.

Burnitt, Christopher, Jared Gars, and Mateusz Stalinski, "Politics of Food: An Experiment on Trust in Expert Regulation and Economic Costs of Political Polarization," 2024. Mimeo.

Chandrasekhar, Arun G, Esther Duflo, Michael Kremer, João F. Pugliese, Jonathan Robinson, and Frank Schilbach, "Blue Spoons: Sparking Communication About Appropriate Technology Use," Working Paper 30423, National Bureau of Economic Research September 2022.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Chopra, Felix, Ingar Haaland, and Christopher Roth, "The demand for news: Accuracy concerns versus belief confirmation motives," *The Economic Journal*, 2024, *134* (661), 1806–1834.

Conlon, John J, Malavika Mani, Gautam Rao, Matthew W Ridley, and Frank Schilbach, "Learning in the Household," Technical Report, National Bureau of Economic Research 2021.

Corno, Lucia, Eliana La Ferrara, and Justine Burns, "Interaction, Stereotypes, and Performance: Evidence from South Africa," *American Economic Review*, December 2022, *112* (12), 3848–3875.

Dahl, Gordon, Andreas Kotsadam, and Dan-Olof Rooth, "Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams," February 2018.

Danz, David, Lise Vesterlund, and Alistair J Wilson, "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 2022, *112* (9), 2851–2883.

DeGroot, Morris H., "Reaching a Consensus," *Journal of the American Statistical Association*, 1974, *69* (345), 118–121.

Druckman, James N and Jeremy Levy, "Affective polarization in the American public," in "Handbook on politics and public opinion," Edward Elgar Publishing, 2022, pp. 257–270.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm, "Refugee Migration and Electoral Outcomes," *The Review of Economic Studies*, October 2019, *86* (5), 2035–2091.

Enos, Ryan D., "Causal effect of intergroup contact on exclusionary attitudes," *Proceedings of the National Academy of Sciences*, March 2014, *111* (10), 3699–3704.

Fang, Ximeng, Sven Heuser, and Lasse S. Stötzer, "How In-Person Conversations Shape Political Polarization: Quasi-Experimental Evidence from a Nationwide Initiative," ECONtribute Discussion Papers Series 270, University of Bonn and University of Cologne, Germany Dec 2023.

Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia, "Listen to her: Gender differences in information diffusion within the household," *Journal of Public Economics*, 2024, *239*, 105213.

Flaxman, Seth, Sharad Goel, and Justin M. Rao, "Filter Bubbles, Echo Chambers, and Online News Consumption," *Public Opinion Quarterly*, 2016, *80*, 298–320.

Garcia-Hombrados, Jorge, Marcel Jansen, Ángel Martínez, Berkay Özcan, Pedro Rey-Biel, and Antonio Roldán-Monés, "Ideological Alignment and Evidence-Based Policy Adoption," 2024.

Gentzkow, Matthew and Jesse M. Shapiro, "Media Bias and Reputation," *Journal of Political Economy*, 2006, *114* (2), 280–316.

— **and** — , "Ideological Segregation Online and Offline *," *The Quarterly Journal of Economics*, 11 2011, *126* (4), 1799–1839.

Golub, Benjamin and Evan Sadler, "Learning in Social Networks," in "The Oxford Handbook of the Economics of Networks," Oxford University Press, 04 2016.

— **and Matthew O. Jackson**, "Naïve Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, February 2010, *2* (1), 112–49.

González-Bailón, Sandra, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M. Guess, Shanto Iyengar, Young Mie Kim, Neil Malhotra, Devra Moehler, Brendan Nyhan, Jennifer Pan, Carlos Velasco Rivera, Jaime Settle, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker, "Asymmetric ideological segregation in exposure to political news on Facebook," *Science*, 2023, *381* (6656), 392–398.

Graeber, Thomas, Shakked Noy, and Christopher Roth, "Lost in Transmission," CESifo Working Paper 10903 2024.

Guess, Andrew, Benjamin Lyons, Brendan Nyhan, and Jason Reifler, "Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think," White Paper, Knight Foundation 2018.

Guess, Andrew M., "Almost Everything in Moderation," *American Journal of Political Science*, 2021, *4* (65), 1007–1022.

Guriev, Sergei, Emeric Henry, Théo Marquis, and Ekaterina Zhuravskaya, "Curtailing false news, amplifying truth," *Amplifying Truth (October 29, 2023)*, 2023.

Guyer, Joshua J, Leandre R Fabrigar, and Thomas I Vaughan-Johnston, "Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion," *Personality and Social Psychology Bulletin*, 2019, *45* (3), 389–405.

Haaland, Ingar and Christopher Roth, "Labor market concerns and support for immigration," *Journal of Public Economics*, 2020, *191*, 104256.

Henry, Emeric, Ekaterina Zhuravskaya, and Sergei Guriev, "Checking and Sharing Alt-Facts," *American Economic Journal: Economic Policy*, 2022, *14* (3), 55–86.

Hobolt, Sara B, Katharina Lawall, and James Tilley, "The polarizing effect of partisan echo chambers," *American Political Science Review*, 2024, *118* (3), 1464–1479.

Hossain, Tanjim and Ryo Okui, "The binarized scoring rule," *Review of Economic Studies*, 2013, *80* (3), 984–1001.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood, "The origins and consequences of affective polarization in the United States," *Annual review of political science*, 2019, *22* (1), 129–146.

Jackson, Matthew O. and Leeat Yariv, "Diffusion of Behavior and Equilibrium Properties in Network Games," *American Economic Review*, May 2007, *97* (2), 92–98.

Jo, Donghee, "Better the devil you know: An online field experiment on news consumption," *documento de trabajo de la Universidad Northeastern, consultado el*, 2017, *20.*

Kashner, Daniel and Mateusz Stalinski, "Preempting polarization: An experiment on opinion formation," *Journal of Public Economics*, 2024, *234*, 105122.

Levy, Ro'ee, "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 2021, *111* (3), 831–870.

Lin, Winston, Donald P. Green, and Alexander Coppock, "Standard Operating Procedures for Don Green's Lab at Columbia," 2016.

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 2021, *111* (6), 1807–1844.

_ , "Has Intergroup Contact Delivered?," 2024. Mimeo.

Morris, Stephen, "Political Correctness," *Journal of Political Economy*, 2001, *109* (2), 231–265.

Mousa, Salma, "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq," *Science*, August 2020, *369* (6505), 866–870.

Nelson, Jacob L. and James G. Webster, "The myth of partisan selective exposure: A portrait of the online political news audience," *Social media+ society*, 2017, *3* (3).

Ortoleva, Pietro and Erik Snowberg, "Overconfidence in political behavior," *American Economic Review*, 2015, *105* (2), 504–535.

Paluck, Elizabeth Levy, Roni Porat, Chelsey S Clark, and Donald P Green, "Prejudice reduction: Progress and challenges," *Annual review of psychology*, 2021, *72*, 533–560.

_ , Seth A. Green, and Donald P. Green, "The contact hypothesis re-evaluated," *Behavioural Public Policy*, November 2019, *3* (2), 129–158.

Perego, Jacopo and Sevgi Yuksel, "Media competition and social disagreement," *Econometrica*, 2022, *90* (1), 223–265.

Pettigrew, Thomas F. and Linda R. Tropp, "A meta-analytic test of intergroup contact theory," *Journal of Personality and Social Psychology*, 2006, *90* (5), 751–783.

Rao, Gautam, "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 2019, *109* (3), 774–809.

Robbett, Andrea, Lily Colón, and Peter Hans Matthews, "Partisan political beliefs and social learning," *Journal of Public Economics*, 2023, *220*, 104834.

Rossiter, Erin L., "The Similar and Distinct Effects of Political and Non-Political Conversation on Affective Polarization," January 2023.

_ **and Taylor N. Carlson**, "Cross-Partisan Conversation Reduced Affective Polarization for Republicans and Democrats Even after the Contentious 2020 Election," *The Journal of Politics*, October 2024, *86* (4), 1608–1612.

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

Scacco, Alexandra and Shana S. Warren, "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria," *American Political Science Review*, August 2018, *112* (3), 654–677.

Schindler, David and Mark Westcott, "Shocking Racial Attitudes: Black G.I.s in Europe," *The Review of Economic Studies*, January 2021, *88* (1), 489–520.

Schwardmann, Peter and Joel Van der Weele, "Deception and self-deception," *Nature human behaviour*, 2019, *3* (10), 1055–1061.

_ **, Egon Tripodi, and Joël J Van der Weele**, "Self-persuasion: Evidence from field experiments at international debating competitions," *American Economic Review*, 2022, *112* (4), 1118–1146.

Smith, Vicki L and Herbert H Clark, "On the course of answering questions," *Journal of memory and language*, 1993, *32* (1), 25–38.

Sunstein, Cass R, *Republic.com*, Princeton university press, 2001.

_ , *Going to extremes: How like minds unite and divide*, Oxford University Press, 2009.

_ , *# Republic: Divided democracy in the age of social media*, Princeton university press, 2017.

Yuksel, Sevgi, "Specialized learning and political polarization," *International Economic Review*, 2022, *63* (1), 457–474.

A. Additional Tables and Figures

Table A.1: Quiz Questions and Answers

(1) #	(2) Label	(3) Question	(4) Options
Q1	Inflation	What was the path of inflation over the last three years?	a) It was low throughout b) It was high throughout c) It first increased and then decreased d) It increased throughout e) It decreased throughout
Q2	Declare War	Which part of government has the power to declare war?	a) Congress b) The Senate c) The President d) The Department of Defense e) The Secretary of State
Q3	Spend Least	What does the US government currently spend the least on?	a) National Security b) Healthcare c) Social Security d) Foreign Aid e) Education
Q4	Enforce Laws	Which branch of government is responsible for carrying out and enforcing laws?	a) The Legislative branch b) The Judicial branch c) The Executive branch d) The Deliberative branch e) The Enforcing branch
Q5	Gun Checks	According to a survey of US gun owners, what percentage of guns is obtained without background checks?	a) Exactly zero percent b) Between 0 and 25 percent c) Between 25 and 50 percent d) Between 50 and 75 percent e) More than 75 percent
Q6	House Speaker	Who among the following was never a speaker of the US House of Representatives?	a) Nancy Pelosi b) Mike Johnson c) Mitch McConnell d) Kevin McCarthy e) John Boehner
Q7	Healthcare	Compared to countries like Colombia, Finland and Italy, how much of their GDP do Americans spend on health care?	a) A quarter as much b) Half as much c) About the same d) Twice as much e) Four times as much
Q8	Gun Deaths	What is the biggest contributor to gun-related deaths in the United States?	a) Suicides b) Hunting accidents c) Police shootings d) Murders e) Shooting range accidents
Q9	Trump Vaccines	Which of the following statements best describes Donald Trump's most recent view of vaccines, such as those for measles and COVID?	a) They tend to be ineffective b) They are dangerous and can cause autism c) They are very important and people should get vaccinated d) We need more research on whether certain vaccines work e) Everybody should be forced to get vaccinated
Q10	Deported Most	Which administration deported the most immigrants?	a) George Bush b) Bill Clinton c) George W. Bush d) Barack Obama e) Donald J. Trump

Q11	Biden Police	Which of the following statements best describes Joe Biden's position on policing and the 'defund the police' movement?	a) Biden supports "defund the police" unequivocally b) Biden supports "defund the police" unequivocally, but has not condemned police violence against African Americans c) Biden does not support "defund the police", but has condemned police violence against African Americans d) Biden does not support "defund the police" and has not condemned police violence against African Americans e) Biden has never commented on police violence or the "defund the police" movement
Q12	Food Stamps	What fraction of the US population is on food stamps?	a) Between 0% and 10% b) Between 11% and 20% c) Between 21% and 30% d) Between 31% and 40% e) Between 41% and 50%
Q13	Senate Majority	Who is the current senate majority leader?	a) John Fetterman b) Mitch McConnell c) Paul Ryan d) Newt Gingrich e) Chuck Schumer
Q14	Immigrants	Roughly, how many unauthorized immigrants resided in the US in 2021 (including those who overstayed their visas)?	a) 100 000 b) 500 000 c) 1 million d) 5 million e) 10 million
B1	Lowest Tax	Which of the following countries has the lowest corporate income tax?	a) Switzerland b) United States c) Germany d) Mexico e) Ireland
B2	Fewest Police	Which of the following countries has by far the fewest police officers per capita?	a) France b) Spain c) Russia d) United States e) Argentina
B3	Obama Sec. of State	Who was Secretary of State during Barack Obama's second term as president (2013-2017)?	a) Al Gore b) Hillary Clinton c) John Kerry d) Joe Biden e) Ash Carter

Notes: The correct options are highlighted in green. The questions and options are presented to the participants in the order displayed above. The three questions denoted "B" in Column (1) are surprise bonus questions that were asked in the post-experiment survey to detect participants that use google or AI to find correct answers.

Table A.2: Attrition

	(1) Attrition in Main Experiment	(2) Attrition in Main Experiment	(3) Completed Follow-Up	(4) Completed Follow-Up
Crosspartisan Interaction	0.0389	0.0353	0.0346	0.0274
	(0.0320)	(0.0323)	(0.0302)	(0.0306)
Player Party: Democrat		-0.0247		-0.0465
		(0.0226)		(0.0298)
Sample mean	0.202	0.202	0.687	0.687
Observations	1245	1245	993	993
R^2	0.002	0.003	0.001	0.004

Notes: Columns (1) and (2) report regression results for attrition among participants who reach the treatment screen and who are not assigned the role of deciders. 1245 participants were assigned to a treatment and not assigned the decider role, and 993 of them successfully completed the main experiment. Columns (3) and (4) show regression results for successfully completing both the main experiment and the follow-up survey, which was administered, on average, 98 days after the main experiment. Of the 993 participants who completed the main experiment, 682 completed the follow-up survey. The regressions are not reweighted. Standard errors clustered at the pair level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Table A.3: Incentivized Expected Improvement Elicitation

	(1) Improvement	(2) Bias
Incentive for E[Improv]	-0.00344	-0.272
	(0.128)	(0.225)
Sample mean	1.447	1.065
Observations	993	993
R^2	0.000	0.002

Notes: This table presents the results of regressing actual improvement and a measure of estimate bias (constructed as the difference between expected improvement prior to the interaction and actual improvement) on an indicator for whether the expected improvement question was incentivized. Recall that, in order to check whether incentives induce participants to game the expected improvement question, we incentivized the expected improvement question only for half of our participants. Participants learned whether an expectation was incentivized after they stated the expectation but before they answered the revised quiz. Republican/Democrat-only pairs are weighted to address sample imbalances. Standard errors clustered at the pair level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Table A.4: Balance robustness

Participants who...	reached random matching,			completed the conversation,			completed follow-up.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Co (weight)	Cross (weight)	p-value (1)-(2)	Co (weight)	Cross (weight)	p-value (4)-(5)	Co (weight)	Cross (weight)	p-value (7)-(8)
Age	42.485	42.48	0.998	42.107	42.581	0.594	44.275	44.716	0.678
	(0.613)	(0.588)		(0.630)	(0.629)		(0.753)	(0.749)	
Female	0.469	0.448	0.492	0.475	0.458	0.589	0.481	0.437	0.266
	(0.022)	(0.021)		(0.023)	(0.023)		(0.029)	(0.027)	
White	0.794	0.778	0.506	0.781	0.784	0.893	0.798	0.801	0.923
	(0.017)	(0.018)		(0.018)	(0.019)		(0.022)	(0.022)	
Race: Black	0.143	0.132	0.608	0.149	0.123	0.242	0.135	0.114	0.402
	(0.015)	(0.015)		(0.016)	(0.015)		(0.018)	(0.017)	
Race: Asian	0.084	0.110	0.138	0.090	0.113	0.224	0.089	0.114	0.274
	(0.012)	(0.013)		(0.013)	(0.014)		(0.016)	(0.017)	
Latino Identity	0.077	0.062	0.36	0.077	0.066	0.49	0.073	0.047	0.15
	(0.012)	(0.010)		(0.012)	(0.011)		(0.015)	(0.011)	
Graduated College	0.220	0.220	0.995	0.205	0.211	0.808	0.206	0.185	0.495
	(0.018)	(0.018)		(0.018)	(0.019)		(0.022)	(0.021)	
Household Income over 50k	0.713	0.721	0.764	0.705	0.715	0.742	0.708	0.704	0.905
	(0.020)	(0.019)		(0.021)	(0.020)		(0.026)	(0.025)	
Urban Residence	0.545	0.517	0.366	0.535	0.520	0.628	0.526	0.487	0.322
	(0.022)	(0.021)		(0.023)	(0.023)		(0.029)	(0.027)	
Republican	0.518	0.495	0.471	0.501	0.499	0.940	0.519	0.513	0.892
	(0.022)	(0.021)		(0.023)	(0.023)		(0.028)	(0.027)	
Voted for Trump	0.430	0.393	0.221	0.412	0.402	0.770	0.439	0.425	0.736
	(0.023)	(0.021)		(0.024)	(0.022)		(0.029)	(0.027)	
Voted for Biden	0.461	0.495	0.253	0.473	0.491	0.578	0.462	0.478	0.689
	(0.022)	(0.021)		(0.023)	(0.023)		(0.028)	(0.027)	
Affective Polarization (baseline)	39.283	39.769	0.781	38.986	39.967	0.597	39.940	39.370	0.803
	(1.249)	(1.222)		(1.318)	(1.306)		(1.671)	(1.564)	
Confidence in Initial Quiz	64.305	63.792	0.597	64.632	63.935	0.490	65.427	64.931	0.679
	(0.687)	(0.683)		(0.718)	(0.708)		(0.886)	(0.809)	
Score in Initial Quiz	6.478	6.552	0.682	6.448	6.583	0.482	6.474	6.900	0.071*
	(0.127)	(0.129)		(0.135)	(0.137)		(0.166)	(0.167)	
Observations	567	545	1,112	510	487	997	341	341	682

Notes: This table reports covariate balance tests for participants who reach three different steps of the experiment, alternative to the attrition results presented in Appendix Table A.2. Each block reports means and robust standard errors in parentheses across treatment groups. Republican/Democrat-only pairs are weighted to address the lower number of Republican-only pairs in the sample. Columns (1) and (2) report the sample of participants who reached the random assignment stage. 1245 participants were randomly assigned to a treatment, and only 1112 of those participants reported all covariates, resulting in a smaller sample. Columns (4) and (5) report the sample of people who participated in a conversation. There are 999 participants who were randomly matched and participated in a conversation. Of these 999 participants, 993 complete the main survey. Two participants are excluded from this sample as they did not report all covariates. Columns (8) and (9) report the sample of people who successfully completed the follow-up survey. Columns (3), (7), and (9) report p-values from a Kruskal-Wallis test, * p<0.1, ** p<0.05, *** p<0.01.

Table A.5: Specification checks (p-values)

	(1) Weights only	(2) Weights and controls	(3) Controls only	(4) No weights, no controls	(5) IP weights only	(6) IP weights and controls
WTP (Fig. 3A)	0.046	0.040	0.031	0.063	0.063	0.029
Negative WTP (Fig. 3B)	0.008	0.007	0.005	0.008	0.008	0.004
Expected improvement (Fig. 5A)	0.001	0.001	0.001	0.003	0.003	0.001
Actual improvement (Fig. 5B)	0.064	0.100	0.087	0.101	0.100	0.076
Feeling thermometer (Fig 9B, top row)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Affective polarization index (Fig 9B, top row)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Feel. therm. at follow-up (Fig 9B, bottom row)	0.009	<0.001	<0.001	<0.001	<0.001	<0.001
Aff. pol. index at follow-up (Fig 9B, bottom row)	0.053	0.005	0.007	0.003	0.002	0.008

Notes: This table shows the p-values for our main results under different regression specifications. Different rows correspond to different outcome variables, with the corresponding figure in the main text enumerated in parentheses. Column (1) shows our main specification, which reweights Republican/Democrat-only pairs to address sample imbalances. Column (2) additionally includes controls. Column (3) includes controls, but omits the reweighting. Column (4) omits both the reweighting and controls. Column (5) estimates the average treatment effect by inverse-probability weighting and using a probit model to predict treatment assignment. Column (6) reports the same model with controls. Control variables include age, gender, partisan affiliation, and baseline polarization. Standard errors are robust for outcomes measured before the interaction and clustered at the pair-level for outcomes measured during or after the interaction.

Table A.6: WTP robustness check

	(1) WTP ($)	(2) WTP ($)	(3) WTP ($)	(4) WTP ($)	(5) WTP negative	(6) WTP negative
Cross-partisan	-0.631**	-0.650**	-0.836***	-0.836***	0.0751***	0.0754***
	(0.316)	(0.316)	(0.282)	(0.281)	(0.0280)	(0.0279)
Controls		✓		✓		✓
Sample	Full	Full	WTP < 16	WTP < 16	Full	Full
Sample mean	12.185	12.185	11.026	11.026	0.250	0.250
Observations	993	993	867	867	993	993
R^2	0.004	0.021	0.011	0.030	0.008	0.029

Notes: This table reports the results of regressing various WTP measures on a cross-partisan treatment indicator. Columns (3) and (4) restrict the sample to participants whose reported willingness to pay was strictly below 16 (the maximum possible value in the survey). The "sample mean" in Columns (5) and (6) refers to the fraction of the sample who reported negative WTP. Control variables include age, gender, partisan affiliation, and baseline polarization. Republican/Democrat-only pairs are weighted to address sample imbalances. Robust standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Table A.7: WTP robustness check, based on open-ended responses

	Disturbance Afraid	Suspicious	Complicated
Cross-partisan	-0.00287	0.00825	-0.00789
	(0.00696)	(0.00755)	(0.0106)
Sample mean	12.185	12.185	12.185
Observations	993	993	993
R^2	0.000	0.001	0.001

Notes: This table reports the results of regressing hand-coded explanations participants gave regarding their WTP decision on a cross-partisan treatment indicator. "Disturbance afraid" means participants indicated being afraid of being disturbed or interrupted (e.g. by a pet or a child) during the conversation. "Suspicious" indicates that participants believed the conversation would not actually happen and that they were being deceived by the experimenters. "Complicated" indicates that participants did not understand. Republican/Democrat-only pairs are weighted to address sample imbalances. Robust standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Table A.8: Transition frequencies, in %

	(1) Co-partisan	(2) Cross-partisan	(3) p-value
Improvement			
Both wrong to both right	1.34	1.19	0.638
Both wrong to one right	4.25	4.65	0.473
One wrong to both right	22.38	19.74	0.029
Worsening			
Both right to both wrong	0.11	0.06	0.505
One right to both wrong	5.43	5.87	0.477
Both right to one wrong	1.11	1.42	0.249
No change			
Both right to both right	22.11	22.18	0.972
Both wrong to both wrong	25.54	24.04	0.302
One right to one right	17.74	20.85	0.011
Observations	7097	6748	
Participants	509	484	

Notes: Columns (1) and (2) report relative frequencies of transition types between the Initial Quiz and the Revised Quiz, with Republican/Democrat-only pairs weighted to address sample imbalances. The unit of observation is the individual-question level, where the transition is measured for each individual using data on the correctness of each question in the Initial Quiz and the Revised Quiz. Column (3) reports p-values that were obtained from each transition type regressed on cross-partisan treatment, with weighted Republican/Democrat-only pairs and standard errors clustered at the pair level.

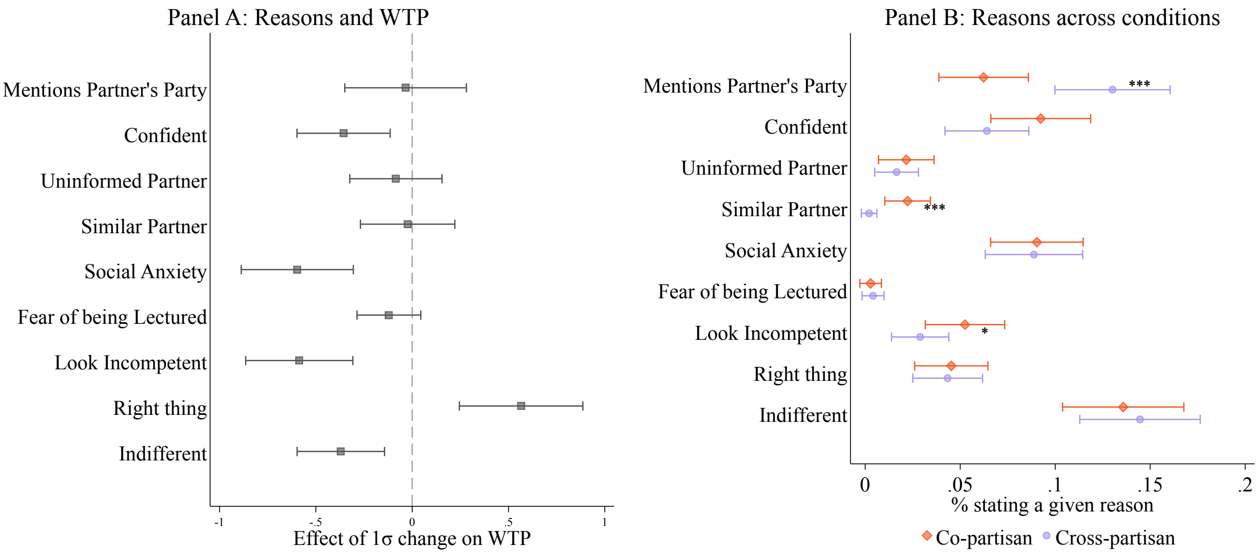

Figure A.1: Reasons for WTP Decision (other)

Notes: Panel A shows the coefficients from regression WTP on hand-coded explanations participants gave regarding their WTP decision. The row variables refer to the following explanations. Mentions Partner's party: Participant mentions the conversation partner's political affiliation; Confident: Participant is confident in their answer and feels that there is nothing to learn; Uninformed partner: Participant feels the conversation partner is unlikely to be well-informed so there is nothing to learn from them. Similar partner: Participant feels the conversation partner will mostly know answers to the same questions as them so there is nothing to learn from them; Social Anxiety: Participant was worried about having the conversation because they are shy or suffer from social anxiety; Fear of being Lectured: Participant does not want to be lectured or confronted; Look Incompetent: Participant is worried that they look incompetent or stupid in the conversation; Right thing: Participant feels that talking to the other person is the right thing to do; Indifferent: Participant mentions that they are indifferent. Panel B shows the predicted values based on regressions of each hand-coded explanation on a cross-partisan interaction indicator. Stars indicate a significant difference between co- and cross-partisan groups. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using robust standard errors from relevant regressions. * p<0.1, ** p<0.05, *** p<0.01.

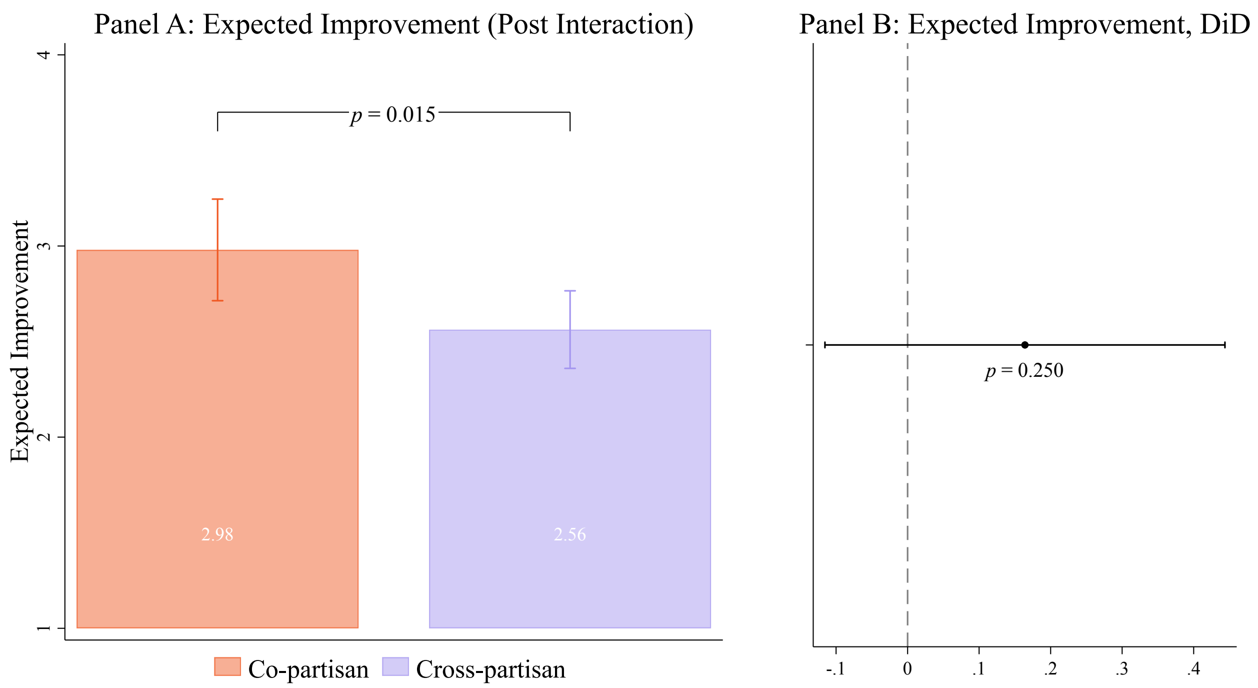

Figure A.2: Expected Improvement (Post Interaction) by Partisan Composition

Notes: Panel A reports the predicted values of regressing expected improvement after the interaction on a cross-partisan treatment indicator. Panel B shows the Difference-in-Difference estimator of expected improvement before and after the intervention. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using standard errors from relevant regressions clustered at the pair level.

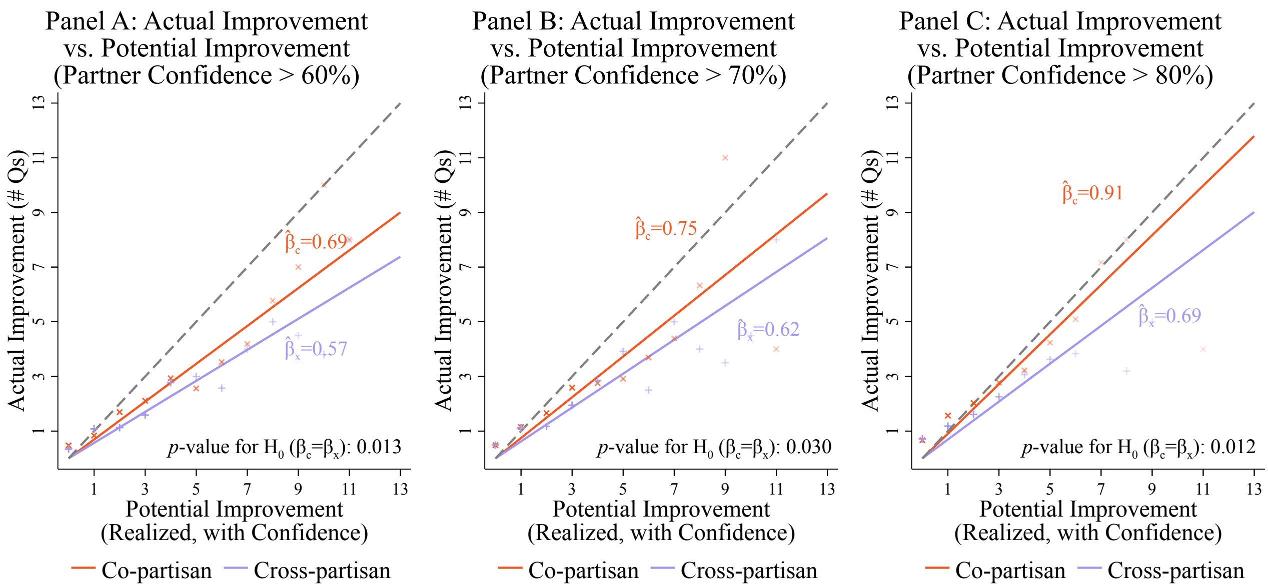

Figure A.3: Actual Improvement and Potential Improvement with Confidence Thresholds

Notes: This figure displays the regression results of actual improvement for each participant on potential improvement (partner confidence > k) across varying confidence thresholds k, following a no-intercept model as specified in Equation 3.1. The measure of potential improvement reflects the number of questions where a participant's incorrect answer could be corrected by their partner's correct answer, conditional on the partner's confidence in their Initial Quiz answer exceeding k. Potential improvement is interacted with treatment assignment in the regression, with β_c and β_x estimating the ease of knowledge extraction under Co-partisan and Cross-partisan treatments, respectively. Panels represent different confidence thresholds for the partner's responses: Panel A (k = 60 percent), Panel B (k = 70 percent), and Panel C (k = 80 percent). In the overlaid scatter plots, Republican/Democrat-only pairs are weighted when calculating the average of actual improvement within each value of the x-axis variable to address sample imbalances; the opacity is weighted by the square root of the number of observations in each bin. We employed the square root of the number of observations in each bin rather than the raw number of observations, because some of the coefficients would have been hardly visible otherwise. In all regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. All outcomes are expressed in terms of number of questions on the quiz (out of 14). The 95 percent confidence intervals and p-values in all panels are computed using standard errors clustered at the pair level.

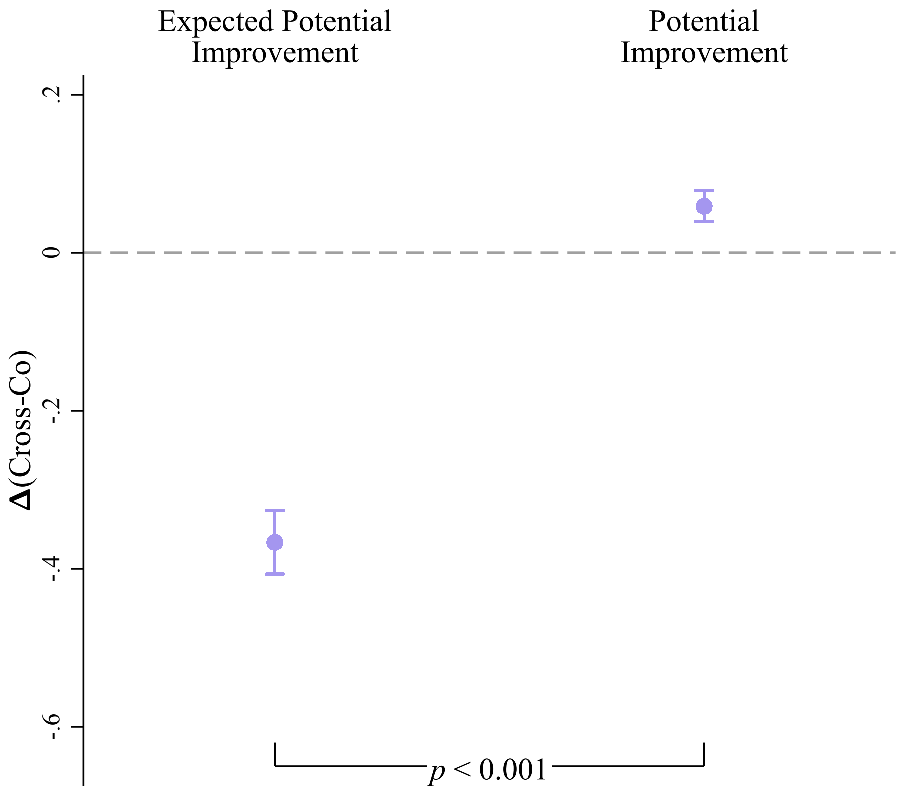

Figure A.4: Expected Potential Improvement and Potential Improvement

Notes: This figure presents treatment effects on expected potential improvement and on potential improvement in number of questions. Expected potential improvement is constructed as follows. First, we consider the beliefs of Democrats from Figure 7. Second, for each question in the quiz, we calculate the expected potential improvement for an affiliate of party $i \in \{Democrat, Republican\}$ who meets with an affiliate of party $j \in \{Democrat, Republican\}$ by taking the probability, according to the beliefs of Democrats, that the affiliate of party i does not know the answer to the question times the probability that the affiliate of party j does. Third, we sum these question-level statistics at the quiz level. This gives us a measure of the degree to which Democrats expect to improve their answers to the quiz when meeting a fellow Democrat or a Republican. Fourth, we repeated the same exercise using the beliefs of Republicans. Fifth, we average the quantities we obtained across cross-partisan and co-partisan interactions. Potential improvement is constructed as follows. First, we select two parties i and j from the set $\{Democrat, Republican\}$. Second, for each question, we calculate the potential improvement for the average affiliate of party i who meets with the average affiliate of party j by taking the sample-level probability that the affiliate of party i does not know the answer to the question times the probability that the affiliate of party j does. These question-level statistics are then summed up at the quiz level and averaged across cross-partisan and co-partisan interactions. The left dot is an estimate of the difference between expected potential improvement in co- and cross-partisan interactions. The right dot is an estimate of the difference between expected potential improvement in co- and cross-partisan interactions. The 95 percent confidence intervals and p-value are computed using bootstrapped standard errors.

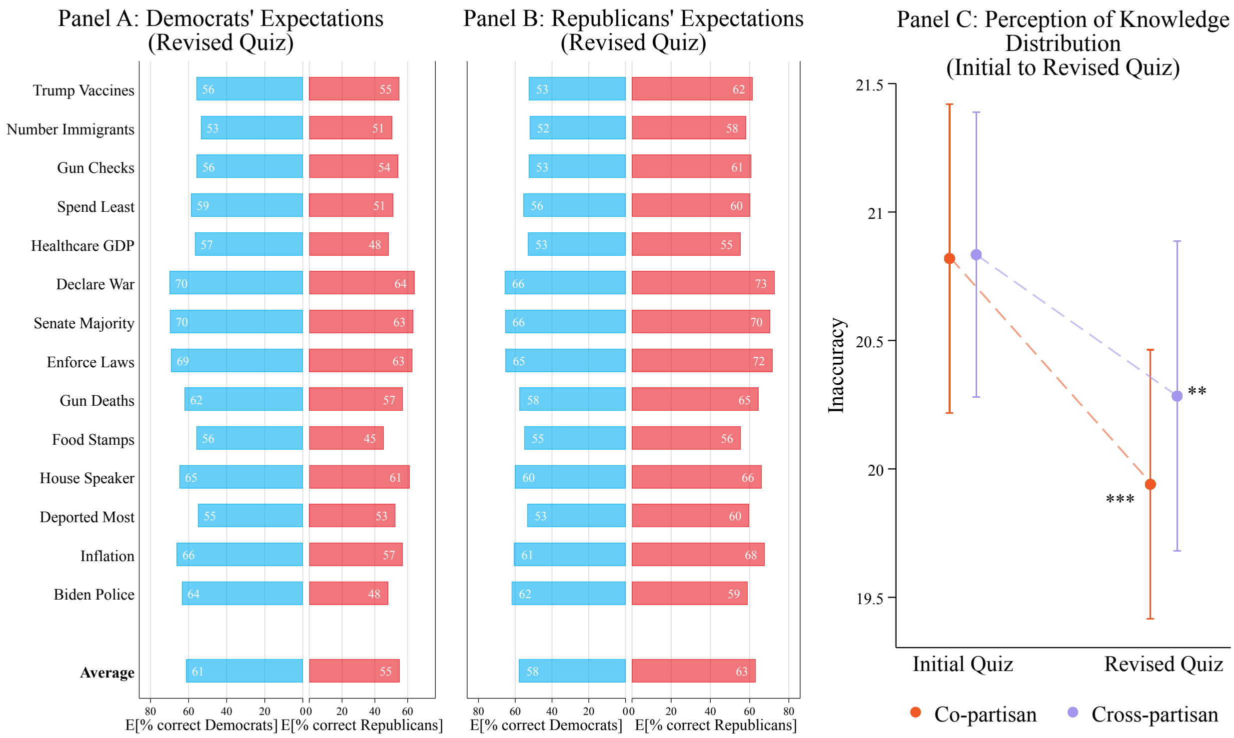

Figure A.5: Expected Knowledge Distribution in Revised Quiz and Forecast Accuracy

Notes: Panels A and B of the figure above present the expected shares of correct responses by respondents affiliated with different parties. In contrast to Figure 7, these expectations are elicited in the context of the Revised rather than the Initial Quiz. Panel A displays the expectations of Democrats about other Democrats and about Republicans. Panel B displays the expectations of Republicans about Democrats and about other Republicans. Panel C benchmarks participants' expectations, both from the Initial and the Revised Quiz, against reality. Specifically, Panel C shows predicted values of the absolute difference between Democrats' and Republicans' beliefs and the actual share of correct answers by Democrats and Republicans. The absolute distance measure is constructed at the individual-question level and is then averaged for each participant. The variable is coded in such a way that a larger absolute difference indicates a greater degree of inaccuracy in the participant's perception of the distribution of knowledge. Panel C also tests whether prediction accuracy improves from the Initial Quiz to the Revised Quiz, and whether it does so differentially for co- and cross-partisan conversations. Prediction accuracy improves in for both co- and cross-partisan pairs, with p-values of 0.000 and 0.036 respectively. There is no evidence that prediction accuracy improves differentially across co- and cross-partisan pairs (p-value 0.351). In the regressions underlying Panel C, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using standard errors clustered at the pair level.

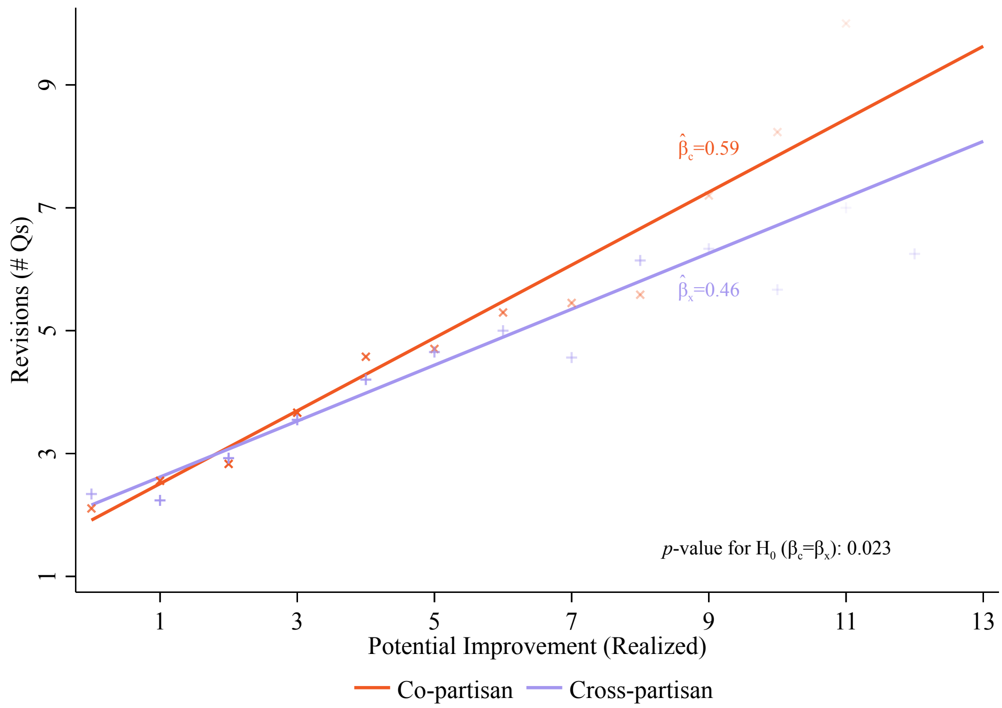

Figure A.6: Number of Revisions and Potential Improvement

Notes: In this figure, we relate potential improvement at the level of individual pairs to the number of revisions in those pairs. We refer to our measure of potential improvement as potential improvement (realized), because it refers to the number of questions, for each participant and for the partner she is randomly matched to in conversation, where the participant gave an incorrect answer and her partner gave the correct one. The figure presents the results of regressing the number of revisions for each participant on potential improvement (realized), in a model where the intercept is fixed at 0. In the regression, potential improvement (realized) is interacted with treatment assignment. β_c and β_x represent the propensity to modify one's answers in Co- and Cross-partisan treatments respectively. The overlayed scatterplot shows the average of the number of revisions within each value of potential improvement; the opacity is weighed by the square root of the number of observations in each bin. We employed the square root of the number of observations in each bin rather than the raw number of observations, because some of the coefficients would have been hardly visible otherwise. In all regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. All outcomes are expressed in terms of number of questions on the quiz (out of 14). The 95 percent confidence intervals and p-value are computed using standard errors clustered at the pair level.

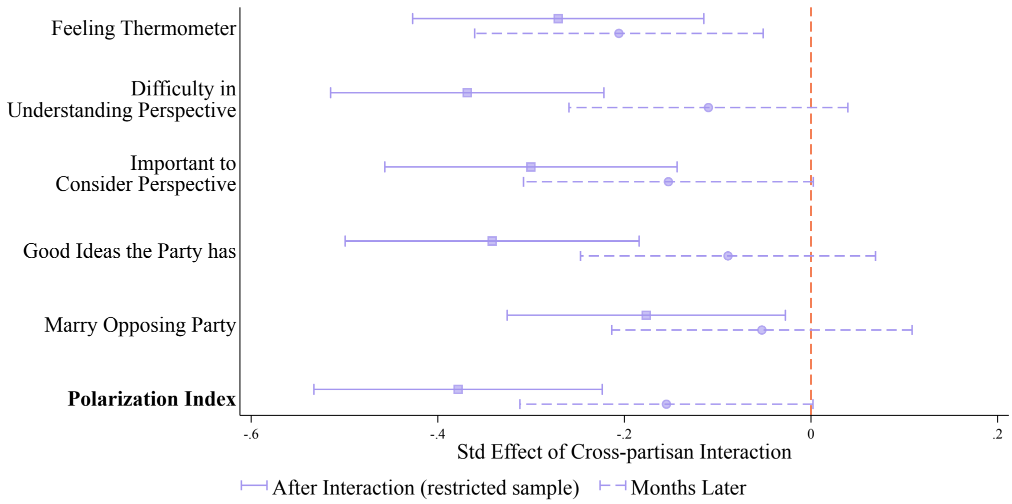

Figure A.7: Affective Polarization (Restricted Sample)

Notes: This figure plots coefficients for the cross-partisan interactions indicator from regressions on the 5 outcome variables and a standardized index of the means of these variables. For all regressions, we restrict the sample to participants who responded to the follow-up survey (N = 682). All outcome variables are standardized by subtracting sample mean and dividing by standard deviation. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using standard errors from relevant regressions clustered at the pair level.

B. Analysis of the Videos of the Conversations

B.1. Data

The audio tracks from each pair's interaction videos were transcribed using the pre-trained transcription model provided by Vosk, an offline open-source speech recognition toolkit. Specifically, we employed vosk-model-en-us-0.42-gigaspeech, a "generic U.S. English model trained by Kaldi on Gigaspeech" (Shmyrev et al., 2024). The transcription output included the recognized words along with their start and end times.

To attribute transcribed words to individual speakers and determine the temporal order of their speeches (a process known as diarization), we used NVIDIA's NeMo framework (Harper et al., 2024). This framework leverages vocal pitch and other audio features to identify speech zones and assign speaker labels. We utilized the pre-configured settings for telephone recordings from the diar_infer_telephonic.yaml file on NeMo's GitHub page and fixed the number of speakers to two. To construct a speaker-level speech dataset, we mapped the diarization output to the transcribed words. [1]

To ensure accuracy, our team manually reviewed each interaction video to match participant identifiers with the speakers recognized by NeMo. We also verified the quality of the diarization output by checking whether speech from different speakers was properly distinguished. [2] Additionally, two research assistants compared all transcripts against the original video calls to resolve any discrepancies in speech recognition.

For categorizing the processed transcriptions, we employed the OpenAI ChatGPT-3.5 Turbo API. The API was used to classify each transcribed segment by determining which quiz question was being discussed or identifying digressions coded as "chitchat." The algorithm analyzed each segment in the context of surrounding text and metadata. To ensure accuracy, two research assistants manually reviewed all categorizations and resolved any errors in the automated classifications.

In addition to transcription, we extracted fundamental pitch (F0) and other acoustic

[1]Occasionally, a single speaker's continuous speech was split into multiple zones. To address this, we merged consecutive segments belonging to the same speaker if the gap between them was less than one second.

[2]In rare cases where the NeMo framework failed to differentiate speakers, we manually corrected the diarization and appended the revised transcription to the dataset.

features from the diarized speech segments using Parselmouth (Jadoul et al., 2018), a Python library for Praat software (Boersma and Weenink, 2021).

The final dataset is structured at the speech-segment level. For instance, a sample transcribed segment might read, “and then there’s the thing it’s deliberative or enforcing.” For analyses conducted at the category level, we concatenated the transcribed text in chronological order.

B.2. Evidence

The conversation data offers valuable insights into how interactions unfold. Figure B.1 examines what participants focus on during their conversations. Although participants were not explicitly required to discuss the quiz, they devoted most of their time to quiz-related topics, with only a small portion of the conversation dedicated to chitchat. Interestingly, the distribution of time across topics shows only minor differences between treatments.

One might hypothesize that participants in cross-partisan interactions might avoid discussing contentious issues where they disagree. If this were the case, such avoidance could contribute to explaining the lower knowledge extraction observed in cross-partisan conversations. Table B.1, however, suggests otherwise: participants spend more time discussing points of disagreement, with this tendency being even more pronounced in cross-partisan interactions. Specifically, column (1) shows that questions involving disagreement between conversation partners are discussed in greater depth, while column (2) reveals that cross-partisan interactions place an even stronger emphasis on disagreements compared to co-partisan ones.

In the next appendix section, we delve further into the conversation data to identify the factors contributing to the treatment gap in knowledge extraction.

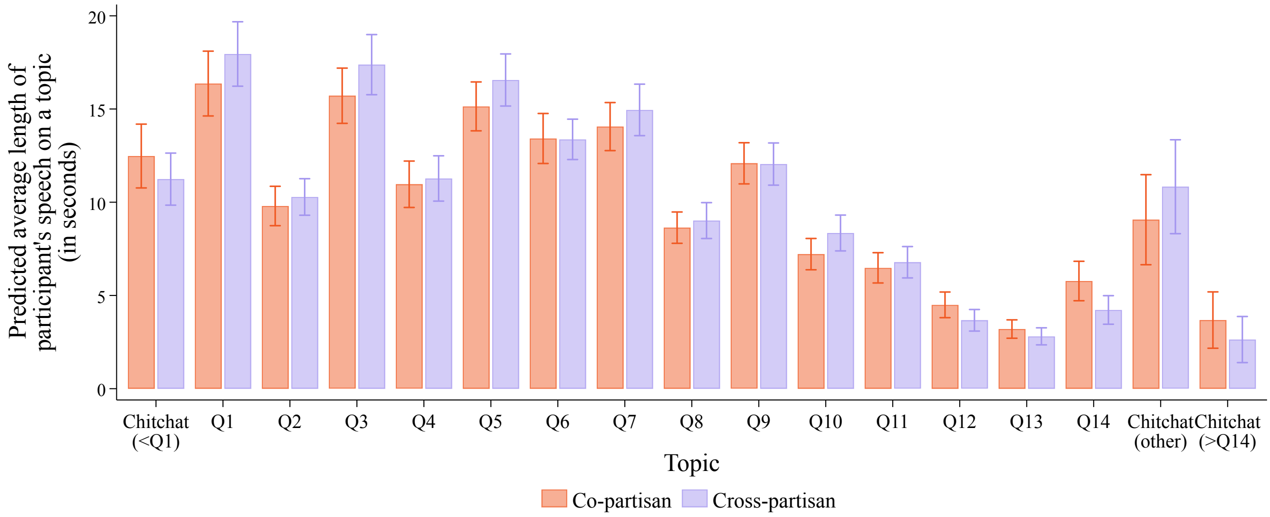

Figure B.1: Speech Length by Question

Notes: This figure illustrates the amount of time participants spend discussing each question and engaging in small talk (chitchat). To identify the questions, denoted in this figure by Q1, Q2, ..., Q14, please refer to Table A.1. The dependent variable is the total duration of a participant's speech within a specific category on the x-axis. The figure presents predicted values from a regression on the participant-question level dataset. The 95 percent confidence intervals are computed using standard errors clustered at the pair level. To address sample imbalances, Republican/Democrat-only pairs are reweighted in the regressions. The speech analysis sample excludes 4 participants: 3 who experienced audio issues and relied on hand gestures to communicate, and 1 whose video was not saved due to a technical glitch.

Table B.1: Share of speech and disagreement

	(1)	(2)
	Nb Words	Nb Words
Disagreement	18.44***	17.00***
	(0.755)	(1.044)
Cross		0.655
		(1.143)
Disagreement × Cross		2.953**
		(1.494)
Social Potency		4.392**
		(1.969)
Confidence on Initial Quiz		-0.00597
		(0.0136)
Question FE	✓	✓
Participant RE	✓	✓
Observations	11192	11192
Participants	984	984

Notes: This table explores the relationship between the amount of time that conversations partners spend on whether they disagree about the answer to that question. Dataset is on participant-question level, and we include random effects on participant level and fixed effects for questions. The speech analysis sample excludes 9 participants: 3 who experienced audio issues and relied on hand gestures to communicate, 1 whose video was not saved due to a technical glitch, and 5 who did not discuss the questions and instead only engaged in small talk (chitchat). Standard errors clustered on pair level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

C. What Explains Differential Extraction of Knowledge?

Table C.1 presents our analysis of the potential drivers behind the differences in knowledge extraction between co-partisan and cross-partisan conversations. In this table, each row in Column 1 reports the treatment effect of being assigned to a cross-partisan conversation on a specific feature of the conversation. Each row in Column 2 is based on a separate regression of actual improvement on potential improvement, the relevant feature of the conversation, and the interaction between potential improvement and that feature. In Section 3.2.4, we defined knowledge extraction as the OLS coefficient of potential improvement in a regression of actual improvement on potential improvement, with the coefficient constrained to be zero. The interaction between a given feature and potential improvement measures whether knowledge extraction is sensitive to that feature; Column 2 displays this interaction effect.

A feature is considered a potential driver of the observed gap in knowledge extraction if it shows a treatment effect (Column 1) and if knowledge extraction is sensitive to it (Column 2). Since this empirical strategy relies partly on correlations, the usual caveats apply, and we consider the evidence presented here as tentative.

The first two rows of Table C.1 pertain to participants' agreement with the statements that their partner knew more than them and that their partner knew what they did not know, respectively. Both variables were measured prior to the conversation. We find that participants are significantly less likely to agree that their partner knew more than them before cross-partisan conversations. However, this qualitative belief does not moderate the rate of knowledge extraction. In contrast, the belief that a partner knew exactly what the participant did not know — that is, the belief in distributed knowledge — does moderate knowledge extraction, but this belief does not exhibit a treatment effect.

The third row examines an incentivized belief, elicited after the conversation, regarding the number of correct answers a participant's partner achieved in the initial quiz. We observe that participants are significantly more pessimistic about their partner's knowledge after cross-partisan conversations, and higher beliefs about the partner's knowledge are associated with greater knowledge extraction.

The fourth row includes a measure of animosity derived from the conversation transcripts using GPT-3.5 Turbo. We used GPT-3.5 to categorize the video transcripts on an animosity scale from -2 (indicating no animosity, friendly conversation, and full mutual respect) to 2 (indicating the opposite). The average animosity score across all participants is -0.27, with a relatively small standard deviation of 0.45. We find that animosity in the conversation is not a significant driver of the observed gap in knowledge extraction.

Rows 5 through 8 focus on qualitative impressions about the partner and the conversation, elicited after the interaction. Row 5 corroborates the finding from row 3 that the belief about a partner's knowledge, elicited post-conversation, is a plausible driver of the treatment gap in knowledge extraction. Row 6 indicates that the belief in distributed knowledge is not a viable candidate. Consistent with the findings in row 4, row 7 shows that cross-partisan partners were not perceived as more confrontational, nor was a partner's confrontational behavior significantly associated with knowledge extraction. Lastly, row 8 reveals that participants in cross-partisan interactions were no less likely to report having had a good time, and that having had a good time did not explain differences in knowledge extraction.

So far, only two rows show both a treatment effect and an interaction term that is statistically significant at the 5 percent level. These rows pertain to measures of trust: specifically, a participant's belief, elicited after the conversation, that their partner was knowledgeable. We interpret this as evidence that a lack of trust in a partner's knowledge during the conversation impeded knowledge extraction. As previously mentioned, we also asked participants, before the conversation, whether they agreed with the statement that their partner "is likely to know more than them" (row 1). However, this prior belief is likely a less accurate measure of trust than the posterior beliefs because it is not informed by the impressions and disagreements that occur during the conversation. This may explain why the prior belief does not account for differences in knowledge extraction.

Rows 9 through 11 are based on the emotions participants reported feeling during the conversation. Using a Likert scale, participants rated the extent to which they felt various emotions. We categorized these emotions into three summary indices, following the influential work of MacKuen et al. (2010): aversion (combining feelings of "bitter," "angry," "disgusted," and "contemptuous"), anxiety ("anxious," "uneasy," and "afraid"), and enthusiasm ("proud," "hopeful," and "enthusiastic"). We find that participants are marginally more likely to report feelings of aversion and anxiety in cross-partisan conversations and significantly less likely to report enthusiasm. However, none of these three emotional indices is associated with higher or lower knowledge extraction.

Rows 12 to 14 pertain to the structure of the conversation. One feature here shows a marginally significant treatment effect and interaction term: the time a participant spent discussing questions where only one of the partners had the correct answer. Time spent on these questions is beneficial for knowledge extraction and is actually higher in cross-partisan pairs; thus, it cannot explain the gap in knowledge extraction. Time spent on questions both partners answered correctly (row 13) and time spent on chitchat (row 14) — which are arguably less productive — also cannot account for the gap in knowledge extraction. The previous Appendix section details how these features were derived and provides a more comprehensive view of their variation across treatments.

Lastly, rows 15 to 17 assess partner characteristics that prior psychological research has found to influence persuasion (Smith and Clark, 1993; Guyer et al., 2019). These include the partner's speech speed, use of filler words, and intonation. However, none of these factors emerge as plausible drivers of the gap in knowledge extraction. Although such features may affect persuasion, they do not appear to influence learning in our context.

Table C.1: Improvement Predictors: Treatment and Interaction Effects

	(1) Treatment Effects	(2) Interaction Effects
1. Before conversation: Partner knows more	-0.176***	0.040
	(0.043)	(0.039)
2. Before conversation: Partner knows what I don't	-0.029	0.082**
	(0.039)	(0.038)
3. Interim belief about partner score	-0.799***	0.026**
	(0.167)	(0.013)
4. Animosity	-0.030	-0.042
	(0.029)	(0.070)
5. After conversation: Partner knew more	-0.136***	0.132***
	(0.043)	(0.036)
6. After conversation: Partner knew what I didn't	-0.029	0.149***
	(0.048)	(0.036)
7. After conversation: Partner was confrontational	0.024	-0.067
	(0.024)	(0.049)
8. After conversation: I had a good time	-0.044	0.005
	(0.044)	(0.043)
9. Aversion	0.014*	-0.242
	(0.008)	(0.198)
10. Anxiety	0.024*	0.128
	(0.013)	(0.172)
11. Enthusiasm	-0.032**	-0.127
	(0.016)	(0.121)
12. Time spent on Qs one has correct	0.032*	0.293*
	(0.017)	(0.153)
13. Time spent on Qs both have right	-0.003	0.386
	(0.010)	(0.268)
14. Time spent chit-chatting	-0.009	-0.346**
	(0.012)	(0.145)
15. Partner: Speed of speech	0.008	-0.007
	(0.035)	(0.017)
16. Partner: No fillers	-0.018	0.031
	(0.015)	(0.023)
17. Partner: Intonation down	-0.006	-0.020
	(0.008)	(0.037)

Notes: In the regressions Republican/Democrat-only pairs are weighted to address sample imbalances. Column 1 shows the results of each row variable regressed on a cross-partisan treatment indicator. Column 2 shows the coefficient of the interaction term of separate regressions of actual improvement on each row variable and the row variable interacted with potential improvement. Standard errors clustered at the pair level in parentheses. * p<0.10, ** p<0.05, *** p<0.010.

D. Predictors of Affective Depolarization in Cross-partisan Conversations

In this appendix, we present the predictors of affective depolarization in cross-partisan conversations. Specifically, Appendix Table D.1 presents several OLS regressions analyzing the change in affective polarization that participants experience from baseline to post-conversation, based on various predictors. The analysis in this appendix is restricted to individuals in cross-partisan conversations.

In column 1, we examine how participants' personal characteristics influence depolarization. We find that being female and, to a lesser extent, younger is associated with greater depolarization. Additionally, participants who exhibit higher baseline affective polarization tend to depolarize more. This is partly due to a mechanical effect, but it also demonstrates that the intervention is not merely "preaching to the choir" and can effectively reach those who are most polarized. Depolarization is also greater among participants with lower social potency, a trait reflecting a tendency toward interpersonal dominance. This finding tentatively supports the hypothesis that such dominance tendencies may hinder warming toward a conversation partner's political party. While Republicans appear to depolarize more than Democrats, this result is not robust across different model specifications (columns 2–6), so we refrain from drawing firm conclusions from it.

Column 2 investigates whether partner characteristics influence depolarization. Being paired with an older partner is associated with greater depolarization, but other partner attributes — such as race, gender, number of correct quiz answers, baseline affective polarization, or social potency — do not significantly predict depolarization. The number of correct answers is plausibly exogenous due to random partner assignment and similar accuracy levels across parties. Conversely, partner age and baseline affective polarization are less plausibly exogenous, given that Republicans in the sample tend to be older and feel less warm toward Democrats than vice versa. Controlling for party affiliation partially mitigates this confounding.

Column 3 focuses on partnership composition, finding that alignment between partners on gender, race, or age does not predict depolarization.

Column 4 examines conversational animosity, derived from video transcripts categorized by GPT-3.5 Turbo on a scale from -2 (no animosity, friendly conversation, mutual respect) to 2 (high animosity). The average animosity score across all participants is -0.27, with a standard deviation of 0.45. Animosity does not significantly predict depolarization.

Column 5 considers participants' self-reported experiences. Those who reported having a good time during the conversation were more likely to depolarize, while perceptions of the partner's knowledge or the conversation's confrontational nature were not predictive. These findings align with earlier results showing that neither a partner's quiz accuracy nor conversational animosity affects depolarization.

Finally, column 6 explores the role of emotions. Participants who experienced feelings of anxiety during the conversation were less likely to depolarize, while those who felt enthusiasm were more likely to do so. The construction of these emotional measures is detailed in Appendix Section C.

In summary, the depolarizing impact of cross-partisan conversations is influenced by individual characteristics (such as gender, age, and social potency), the partner's age, and whether the conversation was enjoyable and evoked feelings of enthusiasm. Partner attributes, including how much knowledge can be gained from them, and homogeneity in partner composition appear to play less significant roles.

Table D.1: Reduction in Affective Polarization from Cross-Partisan Conversations

	(1)	(2)	(3)	(4)	(5)	(6)
Republican	-3.220**	-0.485	0.697	0.660	1.097	1.211
	(1.464)	(1.499)	(1.441)	(1.448)	(1.391)	(1.411)
Age	0.0849*					
	(0.0441)					
Woman	-3.862***					
	(1.349)					
Race: Black	-0.878					
	(2.098)					
Race: Asian	-1.606					
	(2.050)					
Prior affect gap	-0.193***					
	(0.0254)					
Number of questions correct in Initial Quiz	0.337					
	(0.207)					
Social potency	5.431***					
	(2.090)					
Partner: Age		-0.139**				
		(0.0534)				
Partner: Woman		1.406				
		(1.469)				
Partner race: Black		2.821				
		(1.785)				
Partner race: Asian		2.846				
		(2.202)				
Partner: Affective polarization		0.00543				
		(0.0235)				
Partner: Number of questions correct in Initial Quiz		0.159				
		(0.224)				
Partner: Social potency		1.301				
		(2.578)				
Same gender			0.708			
			(1.293)			
Same Race			-0.223			
			(1.352)			
Low age gap			-0.710			
			(1.294)			
Animosity				0.161		
				(1.516)		
After conversation: Partner knew more					-0.685	
					(1.191)	
After conversation: Partner knew what I didn't					-0.826	
					(1.054)	
After conversation: Partner was confrontational					-1.385	
					(1.911)	
After conversation: I had a good time					-2.936***	
					(0.928)	
Aversion						14.89**
						(6.515)
Anxiety						-5.398
						(4.381)
Enthusiasm						-8.262***
						(2.612)
Observations	484	483	483	481	484	484
R^2	0.158	0.029	0.002	0.001	0.026	0.028

Notes: Columns 1-6 report the results of regressing the change in affective polarization, measured as the gap between affect for one's own and affect for the other party at baseline and after a cross-partisan interaction, on various independent variables. The sample for this table consists only of individuals who were part of the cross-partisan treatment, and sampling weights are, unlike in other regressions, not applied here because they would be redundant. Social potency is normalized to vary between 0 and 1. Standard errors in parentheses are clustered at the pair level. p<0.10, ** p<0.05, *** p<0.010.

E. Partisan Differences in Cross-partisan Contact

Our main analyses compared cross-partisan and co-partisan conversations without differentiating between Democrats and Republicans. Here, we revisit the key findings separately for each party. The results are qualitatively similar across groups and align with the aggregated analysis presented in the main text. However, splitting the sample reduces statistical power, limiting our ability to detect partisan differences with confidence.

Figure E.1 shows that Democrats exhibit a significantly lower willingness to pay for cross-partisan conversations compared to co-partisan ones. They are also considerably more likely to express a strict preference against interacting with counter-partisans than against co-partisan interactions. For Republicans, these preferences are less pronounced; the gaps are smaller and not statistically significant.

Both parties expect to learn less from counter-partisans, but this effect is more pronounced and statistically significant only for Republicans (Figure E.2, Panels A and C). When it comes to actual learning, participants from both parties learn less from counter-partisans, but neither effect is statistically significant at conventional levels (Figure E.2, Panels B and D).

Regarding the hedonic experience of the conversation (Figure E.3), both parties enter cross-party interactions with greater pessimism. This pessimism appears stronger among Democrats. Post-conversation, Republicans report enjoying cross-partisan conversations as much as co-partisan ones, while Democrats show a slight and marginally significant preference for co-partisan interactions.

Finally, as illustrated in Figure E.4, cross-partisan conversations lead to substantial reductions in affective polarization for both parties. Since cross-partisan contact requires mutual willingness to engage — given that either participant can choose to walk away — the observed decrease in affective polarization among both Republicans and Democrats underscores the potential of such interactions to foster social cohesion. These findings bolster optimism that policies encouraging cross-partisan contact may promote greater harmony in future exchanges.

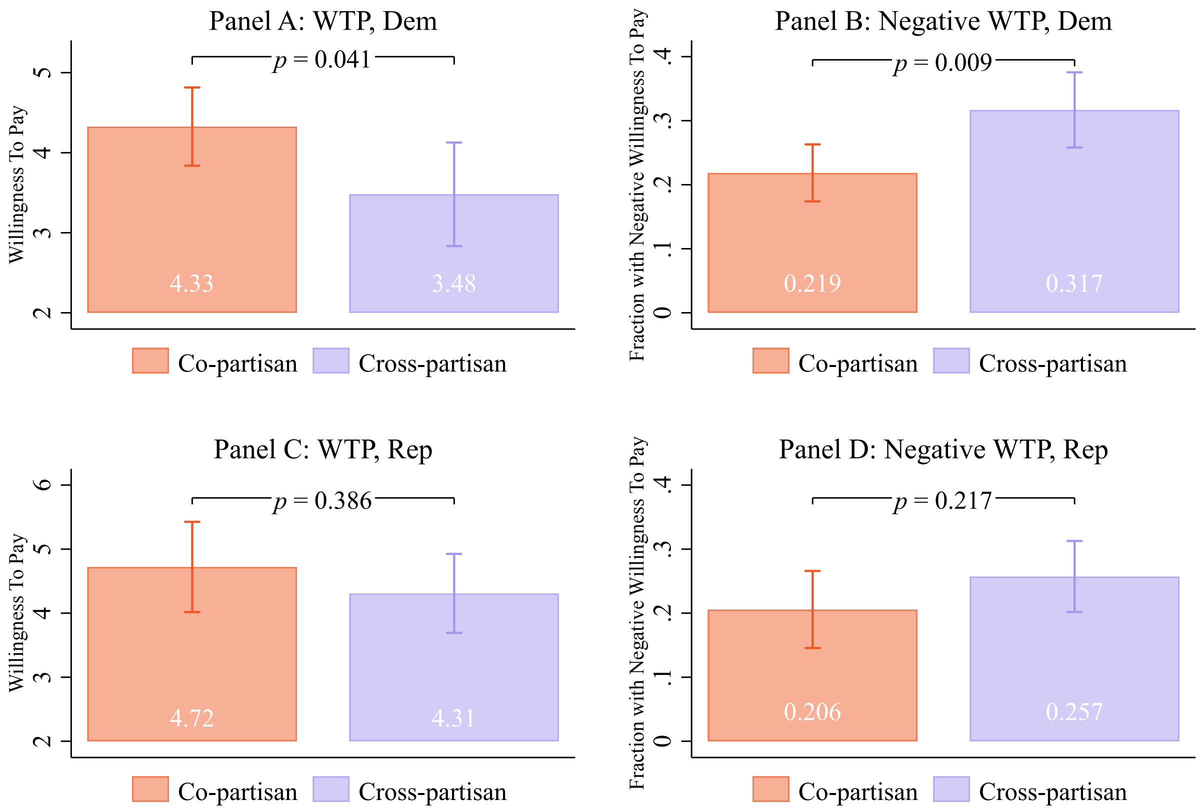

Figure E.1: Willingness to Pay, by Party

Notes: The figure uses predicted values from regressions where Republican/Democrat-only pairs are reweighted to address sample imbalances. Panels A and B report results for Democrats, Panels C and D report results for Republicans. The 95 percent confidence intervals in the figure are computed using robust standard errors from relevant regressions.

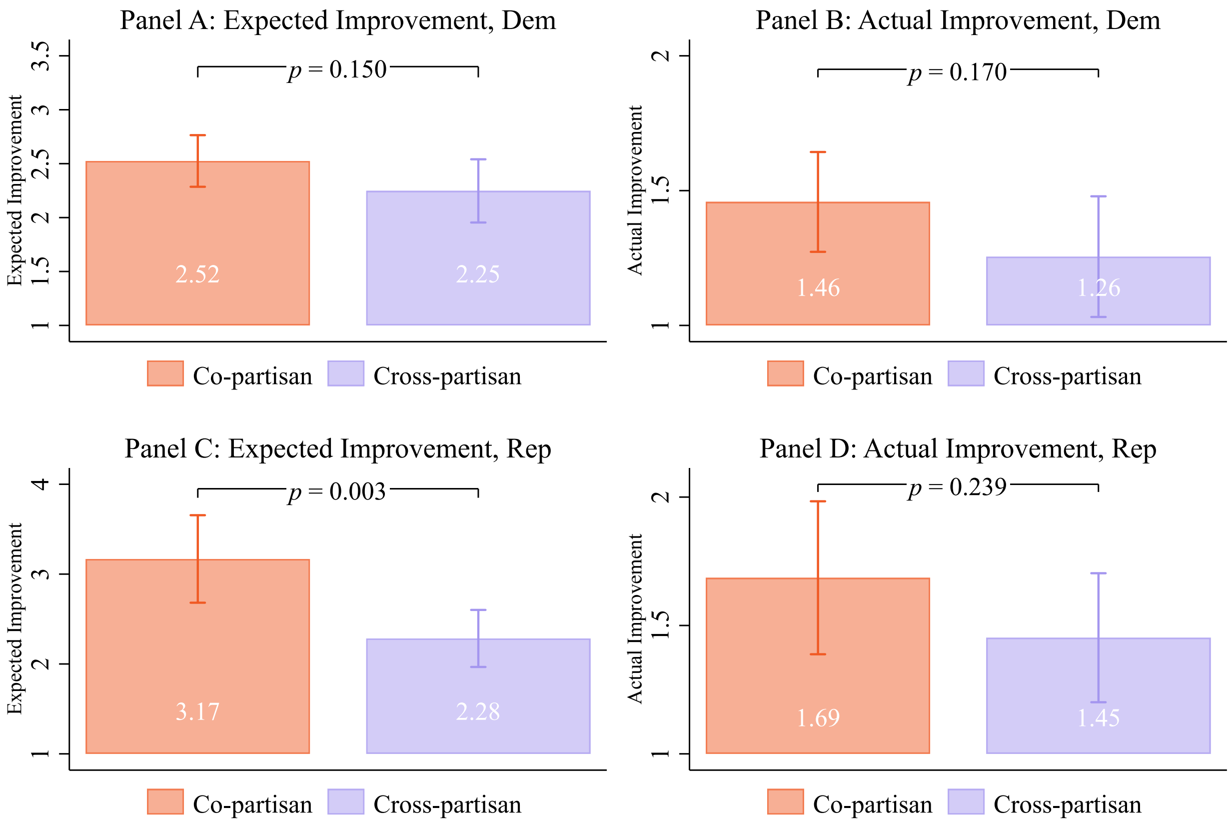

Figure E.2: Expected and Actual Improvement, by Party

Notes: Panel A [C] shows the predicted values of expected improvement regressed on a cross-partisan treatment indicator among our sample of Democrats [Republicans]. Panel B [D] shows the results from regression of actual improvement on a cross-partisan treatment indicator among our sample of Democrats [Republicans]. All panels report predicted values from regressions where Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using robust standard errors for Panels A and C, and standard errors clustered at the pair level for Panels B and D.

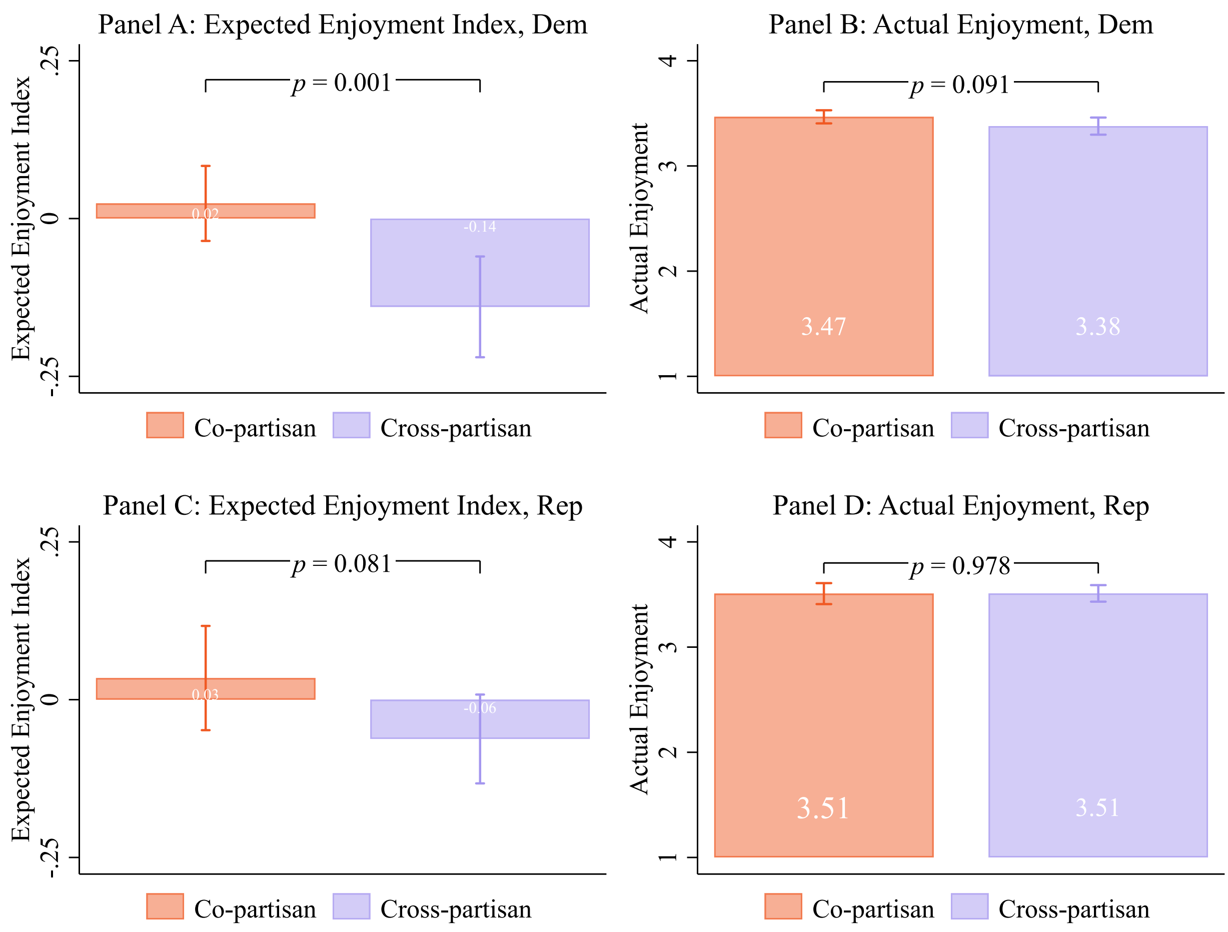

Figure E.3: Expected and Actual Enjoyment, by Party

Notes: Panel A [C] reports the predicted values from a regression of an index of expected enjoyment on a cross-partisan treatment indicator among our sample of Democrats [Republicans]. The enjoyment index is coded on a scale from -1 to 1, with -1 indicating that the participant would not enjoy the interaction and 1 indicating that they would enjoy it. Panel B [D] reports the predicted values of regressing actual enjoyment after the interaction on a cross-partisan treatment indicator among our sample of Democrats [Republicans]. The outcome in Panel B [D] uses a Likert scale from 1 to 4, with 1 indicating "strongly disagree" and 4 indicating "strongly agree." In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals and p-values in the figure are computed using robust standard errors in Panels A and C and using standard errors clustered at the pair level in Panels B and D.

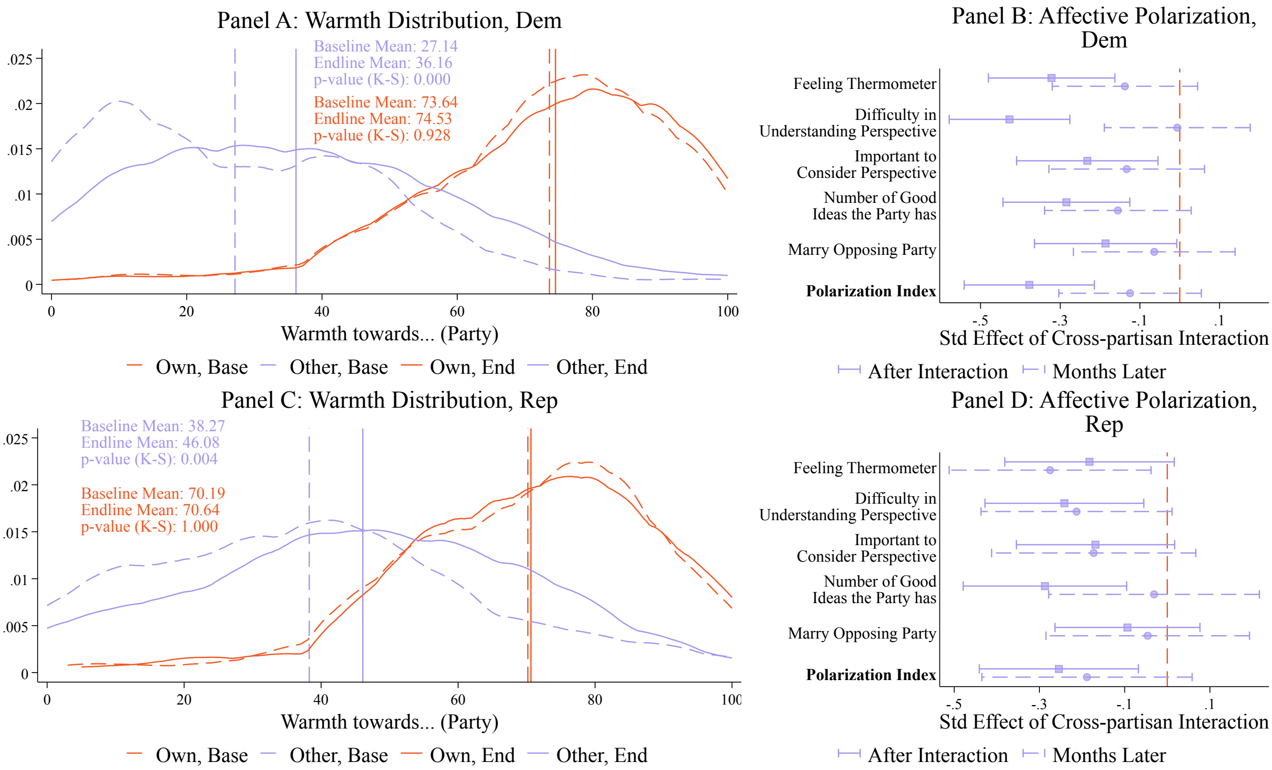

Figure E.4: Affective Polarization, by Party

Notes: Panel A [C] shows the distributions of baseline and endline warmth towards one's own and the other party, restricting the sample to the cross-partisan treatment group and to Democrats [Republicans] only. Panel B [D] plots coefficients on an indicator for a cross-partisan conversation from regressions on the 5 outcome variables along with an equally weighted index of these variables among our sample of Democrats [Republicans]. All these outcomes are measured twice, after the interaction and months later (∼100 days later, on average), and standardized by subtracting the sample mean and dividing by standard deviation. In the pre-registration we indicated that this index would exclude the fifth polarization outcome, already found to be less malleable in Levy et al. (2022). We deviate from the pre-registration, by including this fifth variable in the polarization index, to provide a more comprehensive and conservative picture of our results. In the regressions, Republican/Democrat-only pairs are reweighted to address sample imbalances. The 95 percent confidence intervals in the figure are computed using standard errors from relevant regressions clustered at the pair level.

F. Factual Polarization

F.1. Correlates of Factual Polarization

What drives factual polarization? To explore this question, we construct a measure of baseline factual alignment, which captures the extent to which participants' responses to the baseline quiz align with Democratic or Republican perspectives. This measure ranges from very Democratic (1) to very Republican (5).[3]

Table F.1 shows that baseline factual alignment is significantly correlated with two key factors: self-reported ideological intensity, which ranges from very liberal (-3) to very conservative (3), and news consumption slant, measured as the average slant of the self-reported media consumed, ranging from very liberal (-1) to very conservative (1).

Table F.1: Correlates of Baseline Factual Alignment

	(1) Factual Alignment	(2) Factual Alignment
News Slant	0.156***	
	(0.0597)	
Reported Ideology		0.0268***
		(0.00735)
Observations	417	993
Sample mean	2.538	2.567
R^2	0.015	0.013

Notes: To construct the baseline factual alignment, we use ChatGPT-4 to rank all possible answers to quiz questions on a scale from the most aligned with the Democratic views (1) to the most aligned with Republican views (5). Then, for each participant in the study, we calculate their average alignment at baseline across the 7 quiz questions in which the share of correct responses among Democrats and Republicans is significantly different. The intensity of a participant's ideology varies from very liberal (-3) to very conservative (3). News consumption slant is the average slant of the media that a participant reports consuming. It varies from very liberal (-1) to very conservative (1). Republican/Democrat-only pairs are weighted to address sample imbalances. The lower number of observations in column 1 is due to the fact that not everyone consumes news. Robust standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

F.2. Treatment Effect on Factual Polarization

To analyze the degree of factual polarization among Democrats and Republicans within each treatment condition, we use linear discriminant analysis (LDA). LDA is particularly

[3]Specifically, using ChatGPT-4, we rank all possible answers to the quiz questions on a scale from most aligned with Democratic views (1) to most aligned with Republican views (5). For each participant, we then calculate their average alignment across the seven quiz questions where the share of correct responses differs significantly between the parties.

well-suited for this analysis because it performs well with smaller sample sizes. Specifically, we employ a leave-one-out classification method, predicting each participant's partisan affiliation based on their responses to individual quiz questions. We then calculate the share of participants whose party affiliation is correctly predicted. A higher share of correct predictions indicates that participants' responses are more closely aligned with their party affiliation, suggesting greater factual polarization. To assess statistical significance, we perform chi-square tests for identical distribution across conditions and compute p-values.

Figure F.1 shows that the degree of factual polarization, as proxied by the share of correctly predicted responses, is similar across the two conditions in the initial quiz. However, in the cross-partisan condition, the share of correctly predicted responses decreases significantly, indicating a reduction in factual polarization. Although the differences between conditions become less pronounced in the follow-up survey, they remain statistically significant.

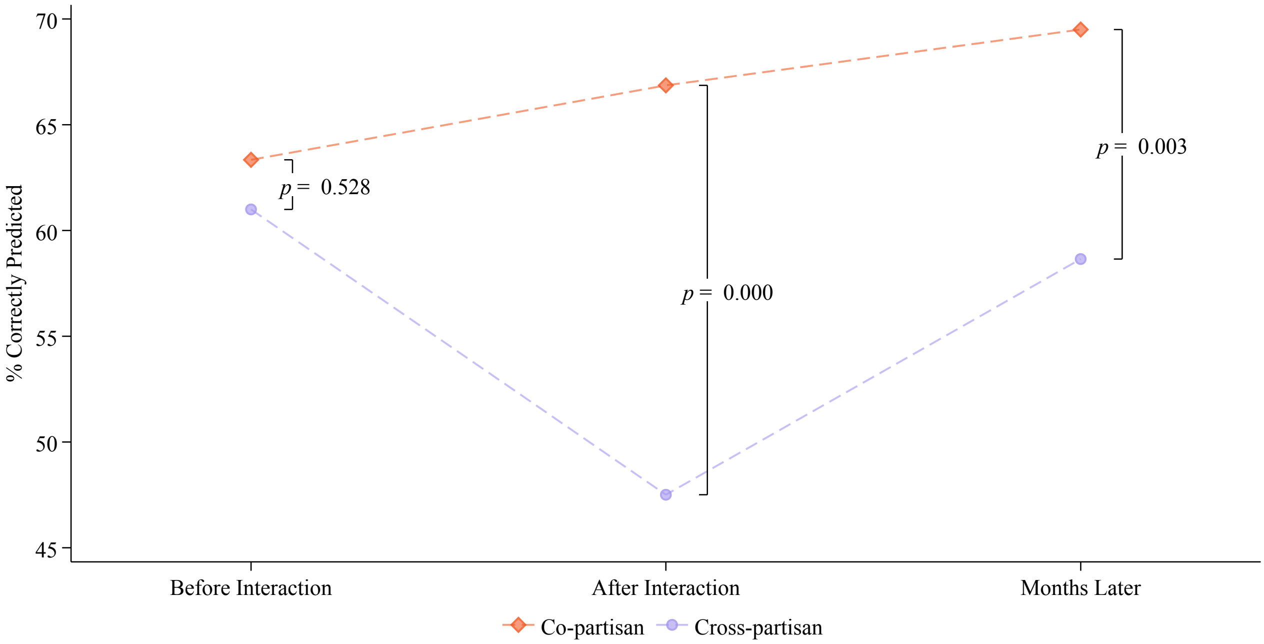

Figure F.1: Factual Polarization, Linear Discriminant Analysis

Notes: We predict the political party affiliation of participants by applying leave-one-out linear discriminant analysis (LDA) to their responses from initial, revised, and follow-up quizzes. The model does not explicitly restrict prediction accuracy to be greater than 50 percent, and accuracy can fluctuate below such "randomness" threshold when no systematic difference between classes is detected. The p-values are derived from Pearson's chi-squared test for whether the distribution of correctly classified observations is the same between the cross-partisan and co-partisan treatments. The analysis before and after the interaction is based on the 993 participants of the main study, and months later only on the 682 participants that completed the follow-up survey. Results are qualitatively the same if we always restrict the analysis to the 682 participants who completed the follow-up survey.

G. Interpersonal Contact: Meta-analysis

We conduct a meta-analysis to examine the short- and long-term effects of interpersonal contact interventions on reducing prejudice and increasing social cohesion. By aggregating evidence, we contextualize our quantitative results within the broader literature and compare design features across studies. This exercise illustrates that our study is the largest to date in estimating the long-term effects of contact interventions and that the effect sizes of cross-partisan contact that we estimate on our pre-registered measures of social cohesion (affective polarization) align with the publication-bias-adjusted meta-analytic effects of this literature (both for short- and for long- term effects).

The studies included in this meta-analysis are detailed in Table G.1. For each study, we identify the outcome of interest and categorize the interventions by type of contact, intensity, mode of interaction (virtual vs. in-person), whether long-term outcomes are measured (and the time frame), and whether beliefs about the expected value of contact were assessed prior to the interaction. Figure G.1 plots the standardized short- and long-term effects of these interventions against study sample sizes. Similar to findings by DellaVigna and Linos (2022) on nudge interventions and Goette and Tripodi (2024) on social recognition interventions, we observe that larger studies tend to report smaller effect sizes—a pattern potentially attributable to publication bias. For both short- and long-term effects, we estimate publication-bias-corrected meta-analytic effects using the methodology of Andrews and Kasy (2019) and report these estimates in the figure.

In Table G.2, we address concerns that specific features of our intervention might limit its comparability to other studies. Notably, the effect size of cross-partisan interactions in our study is comparable to those of studies with similar characteristics (e.g., cross-partisan, pre-registered, and virtual). Moreover, it is not substantially different from the effect sizes observed in other types of interventions.

G.1. Search Terms and Databases

This review encompasses papers published between January 2007 and July 2024. We sourced studies from two key databases: *Scopus* and *Web of Science*. To supplement this collection, we conducted additional searches of bibliographies from related meta-analyses (e.g., Paluck

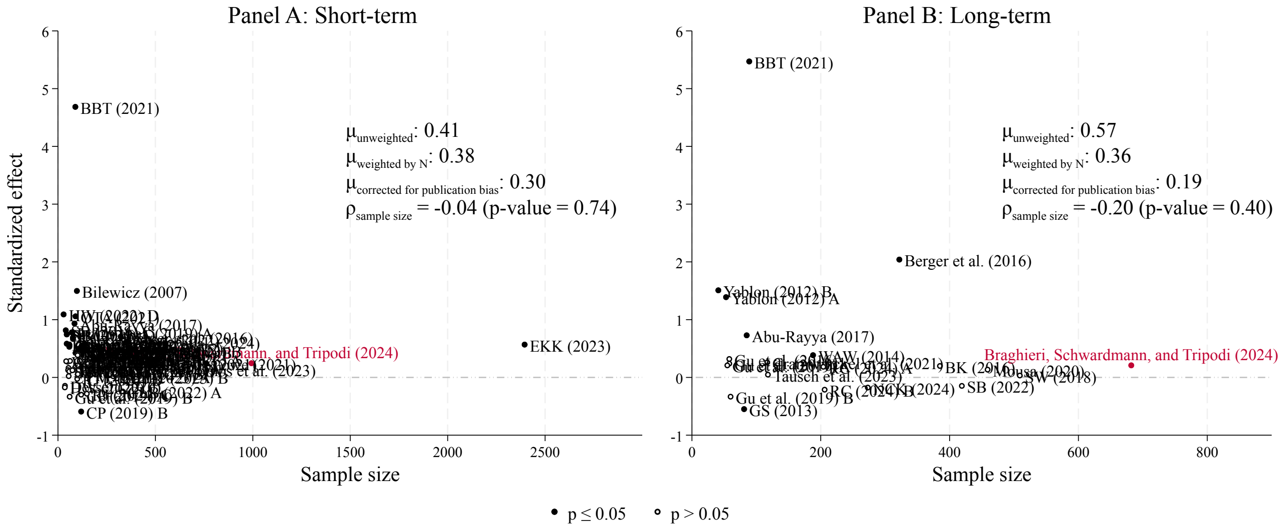

Figure G.1: Meta-analysis

Notes: Each of the effects presented in this chart are described in more detail in Appendix Table G.1. Long-term effect is defined as an effect measured after the date of intervention. Standardized effects are calculated as the ratio between the difference in means and the standard deviation of the control group. The p-values are obtained from t-tests for the equality of means. Estimates corrected for publication bias obtained as in (Andrews and Kasy, 2019), assuming symmetry and t-distribution of effect sizes. When the number of observations are not reported separately for treatment conditions, we impute these by dividing total number of observations by 2. ρ in the scatter plot refers to the correlation between standardized effect and sample size, and p-value refers to the significance of the correlation.

et al., 2021, on prejudice reduction) and performed manual searches on Google Scholar. Additionally, we ensured the inclusion of seminal papers on interpersonal contact published within the specified time frame.

The primary objective of our search was to identify experimental studies aimed at improving social proximity or intergroup attitudes through interpersonal contact. We employed a dual search strategy. The first, a "positive" approach, used terms such as "reconcil*," "social cohesion," "social proximity," "social affinity," "intergroup affinity," "intergroup attitudes," "out-group attitudes," "cross-partisan attitudes," "cross-partisanship," "cross-race attitudes," "intergroup," or ("group" within two words of "tolerance") in proximity to "improv," "enhanc*," or "increas*" (within seven words). The second, a "negative" approach, included terms like "prejudice," "implicit bias," "explicit bias," "ingroup bias," "outgroup bias," "outgroup hate," "intergroup bias," "intergroup anxiety," "affective polarization," "cross-partisan polarization," "political polarization," "echo chambers," "stereotyp*," "racis*," "homophobi*," "islamophobi*," "transphobi*," "ageis*," "abelis*," or "sexis*" in proximity to "reduc*" or "decrease" (also within seven words). Both approaches were paired with qualifying terms such as "quantitative," "study," "experiment*," "controlled study," "intervention," "control group," "dialogue," "interaction," "contact hypothesis," "interpersonal contact," or "contact." These two strategies were combined into unified queries tailored to the structure of each database.

We retrieved papers using these queries for the time range 2007–2024, with retrieval occurring between May and July of 2024.

Scopus

(TITLE-ABS-KEY(((*group W/2 tolerance) OR intergroup OR "outgroup attitude" OR reconcil* OR "social cohesion" OR "social proximity" OR "social affinity" OR "intergroup affinity" OR "intergroup attitudes" OR "intergroup liking" OR "interracial" OR acceptance OR "interpersonal touch" OR "intergroup interactions" OR "out-group attitudes" OR "cross-partisan attitudes" OR "cross-race" OR "racial diversity" OR allies OR outpartisan OR "cross-partisan conversations" OR "Intergroup Dialogue Research" OR "cross-partisanship" OR "cross-race attitudes" OR prejudice OR discrimination OR "Implicit bias" OR "cross-orientation" OR "Explicit bias" OR "In-group bias" OR "outgroup hate" OR "ingroup bias" OR "outgroup bias" OR "intergroup bias" OR "intergroup anxiety" OR "Affective polarisation" OR "Cross-partisan polarisation" OR "Cross-partisan polarization" OR "Political polarisation" OR "Affective polarization" OR "Political polarization" OR "Echo chambers" OR stereotyp* OR racis* OR homophobi* OR islamophobi* OR transphobi* OR ageis* OR abelis* OR sexis* OR xenophobi* OR misogyn* OR "affective polarization" OR "randomized experiments") W/7 (increas* OR enhanc* OR reduc* OR decreas* OR improv* OR better OR alleviate OR minim* OR add OR change OR attenuat* OR develop* OR "long-term" OR bolster* OR effect* OR understand*)) AND TITLE-ABS-KEY((quantitative OR experiment* OR "Controlled study" OR intervention OR "random assignment" OR "randomized experiments" OR contact OR dialogue OR "Control Group" OR interaction OR "Contact Hypothesis" OR "Interpersonal contact" OR “Contact-Based Education” OR “intergroup encounters” OR "intergroup dialogue" OR "intergroup contact" OR "contact theory" OR "intergroup exchange" OR conversation))) AND NOT engineering AND NOT mice AND NOT cats AND NOT animals AND NOT geomagnetic AND NOT neurons AND NOT physics AND NOT nuclear AND NOT cardiac AND NOT metastati* AND NOT surgery AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO(SUBJAREA,"PSYC") OR LIMIT-TO (SUBJAREA,"SOCI") OR LIMIT-TO (SUBJAREA,"NEUR") OR LIMIT-TO (SUBJAREA,"BUSI") OR LIMIT-TO (SUBJAREA,"ECON") OR LIMIT-TO (SUBJAREA,"MULT") OR LIMIT-TO (SUBJAREA,"Undefined")) AND (LIMIT-

TO (LANGUAGE,"English"))

Manual controls:

Year of publication: 2007 – 2024

Web of Science

TS=(((*group NEAR/2 tolerance) OR reconcil* OR "social cohesion" OR "social proximity" OR "social affinity" OR "intergroup affinity" OR "intergroup attitudes" OR "out-group attitudes" OR outgroup OR "cross-partisan attitudes" OR "cross-partisan conversations" OR "cross-partisanship" OR "cross-race attitudes" OR prejudice OR discrimination OR interracial OR "implicit bias" OR "explicit bias" OR "in-group bias" OR "outgroup hate" OR “racial bias” OR “racial diversity” OR “weight bias” OR “outgroup liking” OR “outgroup empathy” OR “dialogue intervention” OR “outgroup hostility” OR "ingroup bias" OR "outgroup bias" OR "intergroup bias" OR "intergroup anxiety" OR "affective polarisation" OR "affective polarization" OR "cross-partisan polarization" OR "cross-partisan polarisation" OR "political polarisation" OR "political polarization" OR "echo chambers" OR racis* OR homophobi* OR islamophobi* OR transphobi* OR ageis* OR abelis* OR sexis* OR “interpersonal liking” OR “intergroup relations” OR “positive out-group attitudes” OR “intergroup encounters”) NEAR/20 (improv* OR enhanc* OR increas* OR reduc* OR decreas* OR attenuat* OR limit OR minim* OR add OR change OR better OR impact OR effect* OR effectiveness OR achiev* OR estimat* OR creat* OR “experimental study” OR vicarious)) AND TS=((quantitative OR experiment* OR "controlled study" OR intervention OR "control group" OR dialogue OR interaction OR "contact hypothesis" OR "interpersonal contact" OR "experimental study" OR "randomized experiment*" OR "random assignment" OR "intergroup contact" OR contact OR interaction OR conversation OR "contact theory" OR “experimental study” OR “intergroup contact” OR “intergroup touch” OR “dialogue intervention” OR “contact-based”)) AND WC=((EDUCATION EDUCATIONAL RESEARCH OR HEALTH CARE SCIENCES SERVICES OR ECONOMICS OR EDUCATION SCIENTIFIC DISCIPLINES OR MULTIDISCIPLINARY SCIENCES OR PSYCHOLOGY SOCIAL OR HISTORY PHILOSOPHY OF SCIENCE OR SOCIAL WORK OR INFORMATION SCIENCE LIBRARY SCIENCE OR HUMANITIES MULTIDISCIPLINARY OR INDUSTRIAL RELATIONS LABOR OR ETHICS OR LANGUAGE LINGUISTICS OR PSYCHOLOGY OR

PSYCHOLOGY APPLIED OR BUSINESS OR FAMILY STUDIES OR NEUROSCIENCES OR ANTHROPOLOGY OR MANAGEMENT OR BEHAVIORAL SCIENCES OR POLITICAL SCIENCE OR PHILOSOPHY OR PSYCHIATRY OR PSYCHOLOGY CLINICAL OR PSYCHOLOGY MULTIDISCIPLINARY OR PSYCHOLOGY DEVELOPMENTAL))
Manual controls:
Year of publication: 2007 – 2024
Language: English
Source type: Article

This search yielded 25,824 studies, from which we subsequently conducted a selection based on the exclusion protocol detailed below, following theoretical and methodological criteria.

Table G.1: Experimental evidence on interpersonal contact and social proximity

(1) Article	(2) Description	(3) Intervention Intensity	(4) Virtual vs. In-Person	(5) Long-Term Outcome	(6) Willingness to Interact with Out-group
Abu-Rayya (2017)	Ethnic contact. Israeli and Ethiopian participants report a significant decrease in intergroup bias ($t = 4.39, p < 0.01, N = 85$) after a collaborative online activity.	3 sessions	Virtual	42	No
Adachi et al. (2015)	Ethnic contact. Negative attitudes of Canadian students towards American students decreased significantly ($t = 3.02, p < 0.01, N = 138$) after a direct-contact video game activity.	One-off interaction	Virtual	—	No
Albuja et al. (2024)	Racial contact. Students randomly assigned to an other-race roommate report significantly higher percentage of other-race friends ($t = 5.06, p < 0.01, N = 356$) afterwards.	One academic year	In-person	—	No
Alimo (2012)	Ethnic contact. Reports more frequent engagement of dialogue participants when compared to waitlisted participants in control group ($t = 5.15, p < 0.01, N = 365$).	One semester	In-person	—	No
Asimovic et al. (2024)	Ethnic contact. Israeli-Jewish and Israeli-Palestinian teenagers report improved intergroup attitudes ($t = 1.77, p = 0.079, N = 138$) after participating in a joint sports activity for a season.	5-8 sessions	In-person	365	No

Baron et al. (2021)	Cross-partisan contact. Undergraduate students in the US reported a decrease in affective polarization ($t = 1.95, p = 0.054, N = 101$) after a couples therapy-style workshop with out-partisans.	One day	In-person	180	No
Benatov et al. (2021)	Ethnic contact. Israeli-Jewish and Israeli-Palestinian students significantly ($t = 26.2, p < 0.01, N = 89$) decreased negative emotional prejudice against the outgroup after playing Minecraft together, both in-person and virtually, with effects lasting at least 6 months after the original intervention.	8 sessions	Both	180	Yes
Berger et al. (2016)	Ethnic contact. Israeli-Jewish and Israeli-Palestinian children participated in direct contact and educational activities, and reported decreased outgroup bias ($t = 6.98, p < 0.01, N = 322$) up to 15 months post-intervention.	12 sessions	In-person	450	Yes
Bilewicz (2007)	Ethnic contact. Jewish participants showed higher levels of out-group liking after meeting young Poles than those who did not ($t = 8.42, p < 0.01, N = 97$).	One-off interaction	In-person	—	No
Bilewicz and Jaworska (2013)	Ethnic contact. After the treatment, Polish and Israeli participants report a more positive attitude towards the outgroup at ($t = 3.94, p < 0.01, N = 137$ and $t = 3.39, p < 0.01, N = 122$)	One-off interaction	In-person	—	No
Boag and Wilson (2014)	Contact with ex-offenders. Criminology students engaged with ex-offenders during a prison visit and reported increased empathy and decreased prejudice ($t = 2.17, p = 0.03, N = 143$) towards prisoners.	One-off interaction	In-person	—	No

Boccanfuso et al. (2021)	LGBT contact. Cisgender participants report significantly lower stigma ($t = 2.83, p < 0.01, N = 114$) towards transgender people following an e-contact intervention.	One-off interaction	Virtual	—	Yes
Broockman and Kalla (2016)	LGBT contact. Cisgender treatment group who interacted with a transgender canvasser report higher acceptance of transgender people ($t = 2.42, p = 0.016, N = 429$) than those in placebo group.	One-off interaction	In-person	90	No
Chang and Peisakhin (2019)	Sectarian contact. Shia and Sunni Muslims watched expert appeal videos and engaged in political discussions. Results from an election game show decreased voting along sectarian lines under no clientelism ($t = 3.15, p < 0.01, N = 119$).	One-off interaction	In-person	—	No
Combs et al. (2023)	Cross-partisan contact. Democrats and Republicans who discussed policy issues in anonymous mobile chats report decreased polarization compared to the placebo group ($t = 1.74, p = 0.08, N = 633$).	One-off interaction	Virtual	—	No
Cook et al. (2019)	Neurodivergent contact. Neurotypical children engaged in a music programme that included contact with autistic children, and showed an increase in prosocial behavior and sympathy with weak statistical significance ($t = 0.88, p = 0.38, N = 49$).	11 sessions	In-person	—	No
Corno et al. (2022)	Racial contact. White students who were randomly allocated a cross-race roommate at a university in South Africa report decreased implicit racial prejudice against Black people ($t = 1.29, p = 0.2, N = 117$).	One academic year	In-person	—	No

Cramwinckel et al. (2021)	LGBT contact. Immediately following direct contact with a sexual minority individual, participants reported decreased negativity towards LGBT inidividuals ($t = 1.22, p = 0.23, N = 117$), thought this effect did not persist at follow-up.	One-off interaction	In-person	9	No
Dessel (2010)	LGBT contact. Public school teachers engaged in dialogues with LGBT minority individuals and reported more positive feelings towards gays with weak statistical significance ($t = 0.28, p = 0.78, N = 36$).	3 laboratory sessions	In-person	—	No
Elwert et al. (2023)	Ethnic contact. A large sample of Hungarian students was randomly assigned to either a Roma or Hungarian deskmate. Participants are likelier to have a Roma friend post-intervention ($t = 8.11, p < 0.01, N = 2395$).	One semester	In-person	—	No
Ermer et al. (2021)	Generational contact. Members of the community participated in performance art and discussions with a senior individual. Participants do not report significant changes in negative stereotypes ($t = 0.083, p = 0.93, N = 55$).	One-off interaction	In-person	—	No
Gabrielli et al. (2022)	Ethnic contact. Italian school children engaged in direct and indirect contact with immigrants alongside an educational program and reported lower implicit prejudice ($t = 2.66, p < 0.01, N = 77$) afterwards.	7 sessions	In-person	—	No
Gaither and Sommers (2013)	Racial contact. White university students who were randomly assigned a cross-race roommate report a significantly higher percentage of non-white friends ($t = 2.95, p < 0.01, N = 140$).	One semester	In-person	180	No

Graham et al. (2014)	LGBT contact. Heterosexual individuals participated in collaborative activities with LGBT confederates and reported more favorable feelings towards LGBT individuals ($t = 2.28, p = 0.024, N = 122$).	One-off interaction	In-person	—	No
Gu et al. (2016)	Urban-rural contact. A collaborative activity bringing together adolescents from rural and urban background led to a decrease in negative outgroup feelings ($t = 2.27, p = 0.027, N = 58$ for Urban, $t = 2.38, p = 0.021, N = 58$ for Rural).	One-off interaction	In-person	300	No
Gu et al. (2019)	Ethnic contact. Direct contact between Malawian shopkeepers and Chinese migrants did not significantly improve outgroup attitudes ($t = 2.42, p = 0.02, N = 58$ for Chinese, $t = 1.18, p = 0.24, N = 60$ for Malawian).	One-off interaction	In-person	10	Yes
Hatoum and White (2022)	LGBT contact. Heterosexual participants reported less intergroup anxiety following an online interaction with a bisexual individual, with male x male dyads showing the greatest effect ($t = 2.94, p < 0.01, N = 29$)	One-off interaction	Virtual	—	No
Hobolt et al. (2024)	Cross-partisan contact. Following a discussion in mixed partisan groups (Conservative and Labor), affective polarization decreased with marginal statistical significance. ($t = 0.18, p = 0.86, N = 283$ for Conservative, $t = 0.087, p = 0.93, N = 375$ for Labor)	One-off interaction	Virtual	—	No
Holoien et al. (2015)	Racial contact. Random assignment of roommates increased the feeling of being understood in Black students who had been paired with White students at weak statistical significance ($t = 0.78, p = 0.44, N = 88$).	10 days	In-person	—	Yes

Influs et al. (2019b)	Ethnic contact. Direct contact between Jewish-Israeli and Israeli-Palestinian teenagers reported increased feelings of empathy towards the outgroup ($t = 3.45, p < 0.01, N = 71$).	8 sessions	In-person	—	No
Influs et al. (2019a)	Ethnic contact. Direct contact between Jewish-Israeli and Israeli-Palestinian teenagers reported increased feelings of empathy towards the outgroup ($t = 0.58, p = 0.56, N = 79$) and also increased oxytocin levels.	8 sessions	In-person	—	No
Kende et al. (2017)	Ethnic contact. Following a friendship-building activity, Hungarians reported an improvement in attitude towards Roma individuals ($t = 0.77, p = 0.44, N = 62$).	One-off interaction	In-person	—	Yes
Koball and Carels (2015)	Overweight contact. Direct contact with an obese individual led to reduced negative feelings in normal-sized people ($t = 3.13, p < 0.01$).	One-off interaction	In-person	—	No
Levendusky and Stecula (2021)	Cross-partisan contact. Republicans and Democrats discussed policy issues in heterogeneous groups and reported significantly lower measures of affective polarization ($t = 3.25, p < 0.01, N = 256$).	One-off interaction	In-person	—	No
Levy et al. (2022)	Ethnic contact. Israeli-Jewish and Arab-Palestinian participants report higher support for intergroup peace-building ($t = 2.22, p = 0.032, N = 45$), with effects lasting up to seven years post-intervention.	8 weeks	In-person	2625	No
Lowe (2021)	Ethnic contact. An intercaste cricket league led to increased positive feelings towards other-caste individuals ($t = 3.15, p < 0.01, N = 762$).	3.5 weeks	In-person	—	No

Lytle and Levy (2015)	LGBT contact. Heterosexual individuals reported a significant improvement in attitudes towards queer individuals ($t = 2.3, p = 0.023, N = 115$) following a confederate who revealed their identity.	One-off interaction	Virtual	—	No
Maunder et al. (2019)	Schizophrenic contact. Following online contact with a schizophrenic person, participants reported lower levels of social distance ($t = 1.99, p = 0.05, N = 90$) towards schizophrenic individuals.	One-off interaction	Virtual	—	No
Mousa (2020)	Religious contact. Christian-Iraqi individuals were randomized to play with either intragroup members or Muslim-Iraqi individuals in a soccer league, though the intervention did not lead to a significant increase in off-the-field out-group positivity ($t = 1.40, p = 0.16, N = 459$).	2 months	In-person	180	No
Nägele et al. (2024)	Refugee contact. Austrian students participated in direct contact with refugees and an educational program. They reported a significant decrease in intergroup anxiety ($t = 3.96, p < 0.01, N = 304$).	One-off interaction	In-person	30	Yes
Ozaydin et al. (2021)	Refugee contact. Nursing students were randomized into internships involving contact with refugee patients and reported a statistically significant reduction in negative feelings towards refugees ($t = 5.67, p < 0.01, N = 90$).	12 weeks	In-person	—	No
Petrík and Popper (2020)	Ethnic contact. Slovak high school students participated in educational programs and direct contact with a Roma individual and reported increased intergroup attitude and trust ($t = 1.34, p = 0.18, N = 150$) following the intervention.	6 sessions	In-person	—	Yes

Rossiter and Carlson (2024)	Cross-partisan contact. Discussions in cross-partisan groups led to a significant decrease in affective polarization and out-party animosity ($t = 3.49, p < 0.01, N = 289$ for Democrats, $t = 0.7, p = 0.48, N = 289$ for Republicans).	One-off interaction	Virtual	3	No
Rossiter (2023)	Cross-partisan contact. Republicans and Democrats engaged in either political or non-political conversations on a chat platform and reported lower affective polarization towards the outgroup ($t = 2.9, p < 0.01, N = 322$ for political, $t = 4.20, p < 0.01, N = 322$ for non-political conversations).	One-off interaction	Virtual	—	No
Rw and Joyce (2023)	Racial contact. White undergraduate students reported lower levels of intergroup anxiety after participating in cross-race counseling sessions ($t = 2.26, p = 0.03, N = 45$ for Asian counselors, $t = 0.094, p = 0.93, N = 99$ for African-American counselors).	One-off interaction	Virtual	—	No
Santoro and Broockman (2022)	Cross-partisan contact. Significant reduction in affective polarization after cross-partisan conversation immediately after ($t = 2.36, p = 0.019, N = 478$), though this effect fades after 3 months.	One-off interaction	Virtual	90	No
Scacco and Warren (2018)	Religious contact. Christian and Muslim men participated in positive intergroup contact and reported higher intergroup cooperation in a dictator game, but no reduction in prejudiced feelings ($t = 0.35, p = 0.73, N = 509$).	16 weeks	In-person	30	No
Stiff and Bowen (2016)	Rival university contact. Across three studies, casually playing a video game with an outgroup member significantly decreased out-group animosity ($t = 2.54, p = 0.015, N = 46$ in study 3).	One-off interaction	In-person	—	No

Tal-Or and Tsfati (2016)	Ethnic contact. Jewish-Israeli and Arab-Palestinian participants watched political films in intergroup company and reported decreased levels of negative out-group attitude ($t = 0.31, p = 0.76, N = 102$).	One-off interaction	In-person	—	No
Tausch et al. (2023)	Cross-partisan contact. Leavers and remainers participated in discussions with outpartisans on issues related to Brexit and reported a decrease in affective polarization ($t = 2.17, p = 0.03, N = 119$).	One-off interaction	Virtual	33	Yes
Voelkel et al. (2021)	Cross-partisan contact. In two experiments, left- and right-leaning students participated in cross-partisan discussions on European policy issues and report an overall reduction in out-party prejudice ($t = 2.47, p = 0.015, N = 107$ for positive political contact).	One-off interaction	Virtual	—	No
Walch et al. (2012)	LGBT contact. Following a panel discussion with a transgender individual, heterosexual participants reported lower levels of transphobia ($t = 0.93, p = 0.36, N = 42$).	2 sessions	In-person	21	No
Welker et al. (2014)	Racial contact. Cross-race contact between White and Black couples led to increase compassion and decreased racial stereotyping ($t = 2.45, p = 0.016, N = 124$).	One-off interaction	In-person	—	Yes
White et al. (2014)	Religious contact. Long-term sustained e-contact between Christians and Muslims decreased affective intergroup bias and anxiety ($t = 2.84, p < 0.01, N = 188$).	9 sessions	Virtual	365	No

Yablon (2012)	Ethnic contact. Direct contact between Israeli-Jewish and Arab-Palestinian individuals increased sentiments of peace-building, but only among those who reported high levels of peaceful motivation prior to the intervention ($t = 4.03, p < 0.01, N = 53$ for Jewish subsample, $t = 4.07, p < 0.01, N = 41$ for Arab subsample).	6 sessions	In-person	7	No
Yücel et al. (2023)	Refugee contact. Direct positive contact between Turkish nationals and Syrian refugees led to increased feelings of empathy and decreased implicit prejudice ($t = 3.13, p < 0.01, N = 207$).	7 days	In-person	—	Yes

Notes: We group intercaste and ethno-religious contact studies with ethnic contact. We further distinguish between racial, ethnic, and refugee contact. The p-values are our own calculations and are obtained from two-tailed t-tests for the equality of means using data provided in the papers, in the replication files, or by the authors. When the original paper does not report the number of observations separately by treatment, we impute these by dividing the total number of observations by the number of treatments. Column 2 reports statistics on the most immediate impacts of the intervention. For studies that report short-term outcomes alone or both short-term and long-term outcomes, these are the short-term results. For studies that only report long-term outcomes, Column 2 instead reflects those. The studies that report only long-term outcomes are: Mousa (2020), Scacco and Warren (2018), White et al. (2014), and Yablon (2012). Column 5 indicates the number of days after the intervention when outcomes are measured, if the intervention includes a long-term follow-up. Column 6 reports whether studies elicit willingness to interact with outgroup members (which we observe to happen primarily in the form of vignettes, or general enjoyment from interacting with the outgroup).

Table G.2: Meta-Analytic Effects, Short-Term

Study sample	N	Unweighted	Publication bias-adjusted
By category			
Ethnic	19	0.71	0.57
Cross-partisan	15	0.28	0.22
LGBT	12	0.29	0.24
Racial	11	0.25	0.22
Refugee	3	0.68	0.53
Miscellaneous	13	0.33	0.10
Pre-registered			
Yes	18	0.30	0.29
No	55	0.46	0.36
Virtual	27	0.32	0.18
In-person	46	0.47	0.40
Full sample	73	0.42	0.30

Notes: This sample consists of 73 estimates from 52 papers that measure short-term effects of contact interventions. Publication bias-adjusted effects obtained as in Andrews and Kasy (2019), assuming symmetry and t-distribution of effect sizes. "Miscellaneous" refers to categories with less than 3 studies and includes neurodivergent, schizophrenic, overweight, religious, generational, sectarian, urban-rural, and ex-offender contact. Studies that include both in-person and virtual contact are categorized as in-person studies in this meta-analysis. Studies that only report long-term outcomes are excluded from this overview.

G.2. Exclusion Protocol

Theoretical grounds for exclusion criteria

Our key concepts of interest are social proximity and interpersonal contact. We hence define clearly our understanding of both terms.

What is *social proximity*? We are interested in the divergence of attitudes and beliefs, associated with prejudice and or animus, among people from opposing groups with little overlap or common ground. We hence understand social proximity as the extent to which these opposing groups share common traits, characteristics, and identities, reflecting this in their behavior (attenuating their prejudice towards the other group).

What is *interpersonal contact*? We include studies that focus on actual and objective interactions, conversational or (positive) physical touch, between two or more individuals from opposing groups.

Given these two definitions, we are particularly interested in studies intending to increase social proximity between groups that are polarized based on a distinct set of identities (e.g., political and religious groups or movements, rival universities, etc) or characteristics (e.g., gender, race, sexual orientation, etc.). In addition, it is paramount that they have at least one outcome assessing prejudice such as attitudes, likability, feeling thermometer, etc. Upon review of associated literatures, we make an explicit exclusion of studies that do not have actual interpersonal contact (e.g., imagined contact studies) and studies that do not explicitly intend to increase social proximity across opposing groups.

Methodological grounds for exclusion criteria

From a methodological perspective, our focus lies on experimental designs and studies based on random assignment allocating participants to treatment and control groups. We also make a clear point to only include studies that have an interpersonal contact intervention. These studies can have either short-term (immediate) and/or long-term post-intervention outcome measures.

Additionally, upon revision of the first studies retrieved, we make certain exclusions based on methodology. Firstly, we exclude any qualitative and non-experimental studies. Secondly, we exclude studies with significant issues of attrition in their sample, since high levels of attrition could affect the unbiasedness of the treatment randomization. It is not uncommon for a study to suffer attrition, so we only avoid studies in which the experimental conditions at the analysis level are severely affected by attrition. A third case for exclusion is that of studies that use cluster-randomized experiments. Such designs are prone to caveats such as having a sharp increase in sampling variance and the possibility of having biased estimators if the design only includes a small number of clusters of unequal size. Finally, we make an exclusion of studies that do not have a clear reporting of results or do not explain their analysis in sufficient detail. In some cases, studies do not fully disclose results on mean effect, standard errors, or sample allocation to treatment and control groups. In such cases, we have to exclude these studies from the review. In other cases, we find that studies do not present results for the entire pooled sample. In these cases, when provided with enough information, we reconstruct the average treatment effect and standard errors (see Table G.1 for pooling of results).

G.3. Coding and Reliability

Following our search and exclusion protocol, we gathered 79 outcomes from 57 studies for the analysis. In such cases, we coded for key characteristics that indicated the papers' contributions to the field of social proximity and interpersonal contact. These characteristics are as follows:

(a) *Type of intervention*. The studies cover a wide range of intervention types within the interpersonal contact category. These include inter-group dialogue, roommate allocations, video discussions, course administration, etc.

(b) *Intensity of intervention*. We distinguish specifically between one-off interactions and repeated contact sessions. This is captured by descriptive labels that are tailored to the study.

(c) *Type of prejudice/polarization targeted*. Studies address different types of factors that affect social proximity (e.g., race, political orientation, nationality, etc.)

(d) *Type of experiment*. We code for the type of experimental design in the study: laboratory, field, or online.

(e) *Primary outcome*. Many studies report on several outcomes to test for multiple hypotheses. We select the primary outcome of interest based on the objective of this review. These outcomes must be related to reducing prejudice or increasing social proximity (i.e., inter-group attitudes, likability, feeling thermometer, etc.) If we find there are several outcomes of interest, these are added to the analysis as an additional hypothesis.

(f) *Relevant experimental conditions and statistical contrasts*. Some studies have key differences in their experimental designs and outcome measurements to capture their hypotheses. One of the most notable distinctions we code is that of the direction of the outcome towards less prejudice or more social proximity.

(g) *Statistics*. We code sample size, mean effects, and standard deviations of treatment and control groups.

(h) *Time of measurement*. We cover studies that have immediate and/or long-term measures of the primary outcomes. We thus code the time interval (in days) between the intervention and the measurement.

(i) *Open science practices*. We are interested in whether the studies were pre-registered or have open data.

G.4. Final Sample

The search culminates in 79 outcomes in 57 studies from the original set of 25,824 studies, after applying the exclusion protocol and the coding procedure.

G.5. Publication Bias Concerns

It is not uncommon to have concerns about the tendency of journals to favor studies that have large and statistically significant effects. We find this is present in our search since there is a strong positive relationship between the effect size and standard errors. We address these concerns by applying publication-bias adjustments to estimated meta-analytic effects (Andrews and Kasy, 2019). We also code for each study whether they include a pre-registration that specifies hypotheses and sample size. We then estimate separate meta-analytic effects for these two groups of studies separately.

H. Instructions: Recruitment Survey

Demographics

Before we start with the main part of the survey, let us ask you a few questions.

What is your year of birth?

[Year]

In which state do you currently reside?

[Dropdown list of possible states]

Generally speaking, do you usually think of yourself as:

[Democrat; Republican; Independent]

Not selected this time (Independents and some Democrats)

Unfortunately, you were not selected to be invited for today's session, but you may be able to do this invitation survey again in the coming days. Please do not return this task, because we intend to approve and pay you for your work nonetheless. Thank you for expressing your interest in participating in our study.

Introduction

Welcome!

This study requires participants to interact. In order to have enough participants in the experiment at the same time, the study will run at a pre-specified date and time. Specifically, the experiment will run **today [Date and time]**.

Here is some additional information about the study:

- The study takes around **25 minutes** to complete.
- The **baseline payment** is **8 pounds**, and you can earn a **bonus payment on top** of the baseline payment of **up to 13 pounds**.
- All incentives in this study are denominated in Pound Sterling (GBP, £). Payments will be converted to and processed in United States Dollars (USD, $) based on the current exchange rate. As of 30 January 2024, 1 GBP is equivalent to 1.27 USD.
- In the study, you may also be asked to have a **brief chat with another participant**. The chat will happen on the screen and with your computer's **webcam** and **microphone** on. The only information we will share between you and the person you meet is your respective first names and political orientations. The person you meet is another participant on [Platform]. Chat recordings will be collected for quality purposes and erased after review.
- **Technical Requirements: This study must be run on a standard web browser on a laptop or desktop. You are required to have a working camera and microphone. Mobile devices are not supported.**

All incentives in this study are denominated in Pound Sterling (GBP, £). Payments will be converted to and processed in United States Dollars (USD, $) based on the current exchange rate. As of 30 January 2024, 1 GBP is equivalent to 1.27 USD.

Note that there will be no deception in the instructions. Everything we tell you about the tasks you face will be implemented in the exact way we tell you. Any analysis and publication will only use data in anonymous form. This study was cleared by the ethics committee of the Hertie School of Governance, in Berlin.

If you experience a technical error or problem, then do not try to restart or retake the study. Rather, send us an email with a description of your problem and we will get back to you. For any questions or complaints, please contact Egon Tripodi (tripodi@hertie-school.org). For interactions with your conversation partner, we discourage you from sharing any personal information beyond your first name, we invite you to be kind, and we ask you to let us know should you need to drop out of the study because of harassment. At any point, even after completing the study, you can contact us to withdraw your consent and to have your data erased.

Before proceeding: Did you already take part in this study in the past couple of months through [Platform]?

[I have done this study before; I have NOT done this study before]

Previous participants screen out

Unfortunately, you cannot complete this study multiple times, but you may participate in other studies conducted by the research team in the future. Thank you for expressing your interest in participating in our study.

Participation

Next you will decide whether you are willing to participate in the study that runs today [Date and time]. Here, you can either choose to earn a bonus of 5 cents now and not participate, or you can give up this bonus to participate in this highly paid study.

If you agree to participate, then **we will send you the study link by Prolific messages and you will have to click the link at [Time]**.

Please make your choice below.

[Yes, I agree to participate in the study that runs today at [Time] and give up the 5 cents bonus; No, I do not agree to participate and will receive the 5 cents bonus]

If agreed to participate

Please enter your participant ID. We will send you a link to the experiment today at [Time] (note the difference with your time zone, if any) via a message on [Platform]. Note that you will be asked to provide your participant ID once again later, when logging into the study.

Please note that this response should auto-fill with the correct ID.

[Field for ID]

If declined to participate

You said you cannot participate this time. Would you like to be invited if we have more sessions of this study?

[Yes; No]

If interested in participating in the future

When would be a good time for you? (Please indicate typical days and times of the week that work for you):

[Open text]

I. Instructions: Main Study

Introduction

Welcome to this study! All of your responses are anonymous, and we have no way of linking them back to your identity. In this study, you will be asked to answer a series of factual questions. These questions were selected based on a group of respondents evenly split between Republicans and Democrats deeming them central to an important and contentious policy debate. Naturally, people differ strongly in their political attitudes. We as researchers do not take a stand on any of the topics we are covering. For this task, it is important that you do not do any research or cheat (a cheating-detection technology is in place). We are only interested in what you currently know.

You will also be asked to have a brief conversation with another participant about these questions. The conversation will happen on the screen and with your computer's webcam and microphone on. The conversation will be recorded to allow quality review. In compliance with data protection regulation (GDPR), recordings are stored separately from the rest of your responses and will be permanently deleted following review from our team.

The only information we will share between you and the person you meet is your respective first names and political orientations. The person you meet is another participant on [Platform].

Technical Requirements: This experiments must be run on a standard web browser on a laptop or desktop. You are required to have a working camera and microphone. Mobile devices are not supported.

All incentives in this study are denominated in Pound Sterling (GBP, £).

[Shown to Connect participants: Payments will be converted to and processed in United States Dollars (USD, $) based on the current exchange rate. As of 30 January 2024, 1 GBP is equivalent to 1.27 USD.]

[Next Button]

Survey

Before we begin, please answer the following three questions.

Are you male or female?

[Male; Female; Other]

What is your first name? (Remember this will only be shared with one person.)

[Open text]

Generally speaking, do you usually think of yourself as:

[Democrat; Republican; Independent]

We would like to get your feelings toward specific groups of people. We will show the name of a group and we would like you to rate the group using something we call the feeling thermometer.

Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 degrees and 50 degrees mean that you do not feel favorable toward the group and that you do not care too much for that group. You would rate the group at the 50 degree mark if you do not feel particularly warm or cold toward the group.

How would you rate **Democrats**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

How would you rate **Republicans**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

[Next Button]

INITIAL QUIZ

Timer

Time left to complete this page: [Countdown from 10 minutes]

You can take up to 10 minutes to complete this quiz.

For each of the fourteen questions below, please determine the correct option. If this initial quiz is selected for payment, then you will receive 0.5£ for each correct answer. Whether the initial quiz determines your payments is randomly determined by the computer and you will learn which mode you are in by the end of the experiment. With a probability of 50 percent the computer puts you in Initial Quiz Mode, which makes the initial quiz count for payment.

For each question below, please also state how confident you are in your answer, i.e. how likely it is that your answer is correct. For example, if you are very uncertain, then you might state a likelihood that your answer is correct between 20 and 30 percent. If you are very certain, then you may pick a likelihood above 90 percent.

Finally, for each question, please state the share of Republicans and the share of Democrats from a previous survey that you think answered the question correctly. You will receive 0.1£ if you are within 5 percentage points of the correct share for one randomly selected question.

You can only proceed if all answers are given and confidence levels are between 20 and 100.

Question	**Your Answer**	**Your Confidence Level**	**Correct Republicans**	**Correct Democrats**
[Question 1]	*[Options 1]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
[Question 2]	*[Options 2]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
		⋮		
[Question 14]	*[Options 14]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*

[Next Button]

CAMERA AND MICROPHONE CHECK

You now have time to test the video chat without a partner. To continue, please allow camera and microphone access in your browser and test that both work. This is a requirement to participate in this experiment. Please click 'Next' at the bottom of the page to continue with the experiment after you have finished testing.

During the video chat, please make sure that you are visible on your camera and can be heard on you microphone. You will not be paid if we can verify that you were not visible or audible.

Test your microphone

Click the button below to start and stop recording. This allows you to test your microphone and speaker.

[Start Test Recording Button that allows recording oneself and playing it back]

Test your camera

If you failed to allow access, then reload the page and the browser might ask you again to allow access. Otherwise, click on the red camera button in the video frame below to receive instructions on how to allow access in your browser manually.

[Video frame that shows oneself]

Do you confirm that your microphone and camera are working correctly?

[Yes; No]

Shown if No

Please contact us using the chat box for further assistance.

Shown if Yes

Thank you for confirming. You may proceed by clicking Next.

[Next Button]

A CHANCE TO INTERACT

In the next part of the experiment, you will be paired with a randomly-drawn participant in the study who took the same quiz as you. You and your partner will have the opportunity to have a 8-minute conversation. The idea is that you discuss your respective answers in the quiz.

After the conversation, you will have a chance to revise your answers to the quiz. Whether the answers you just gave or your revised answers will count towards your final payment is randomly determined by the computer. With a probability of 50 percent the computer puts you in Initial Quiz Mode, which makes the initial quiz count. Otherwise you are put in Revised Quiz Mode, which makes your revised quiz count.

Because this part of the experiment is interactive, and we do not want your partner to have to wait unnecessarily, participation in the experiment requires you to focus exclusively on this study and to complete each page carefully in the time that you are given. Please adhere to the timers and finish the page within the time allocated to receive your payment (between £8 and £21).

[Next Button]

Matching Wait Page

You are being matched with another participant of this study. The study will continue automatically as soon as you are matched. The maximum wait time is 10 minutes. If you do not stay on this tab actively (keep the browser open without using other windows), you cannot be matched and, therefore, you will not paid. If you stay on this tab actively but no one can be found to match you, you will still receive the £8.00 participation fee.

You have been matched!

Timer

Time left to complete this page: [Countdown from 2 minutes]

The computer has now selected a partner for you. Your partner is called [Partner's name] and is a [Partner's party]. Note that, just like you, this person did not choose to interact with you; the two of you were matched at random.

[Next Button]

MAKE A GUESS

Timer

Time left to complete this page: [Countdown from 5 minutes]

Next, you will have to make a guess.

You will have the opportunity to talk to [Partner's name], who is a [Partner's party], and revise your quiz. We will now ask you to give us your best guess as to how many more correct answers you expect to have in the revised quiz compared to the initial quiz.

You may win an additional £0.50 bonus for your guess. Of course, many different outcomes are possible, depending on how knowledgeable [Partner's name] turns out to be and on how the conversation goes. You maximize your chance of winning a £0.50 bonus by simply stating your best guess of how much you will improve on average. The mechanism we use to determine whether you win the bonus is a little complicated, but ensures that you have a higher chance of winning £0.50 the closer your guess about the average improvement is to the true improvement.

Explanation of the mechanism in more detail (optional)

[Initially hidden, shown when clicked]

How your answer to this question maps into your chance of winning the bonus is based on a formula. This formula is designed to make sure that you maximize your chance of winning if you report your true belief of your average improvement on the quiz (therefore, you will likely need to use decimals).

The variable R is the actual improvement, comparing your correct answers on the revised quiz to the original quiz after the experiment. The variable r is what you report for your expectation of the average improvement on the question below. The winning probability for the bonus is then given by winning probability $= 100 - 100 \times (\frac{R-r}{28})^2$.

This probability is higher, the closer is your report to the actual improvement.

Please consider the following points carefully, as they speak to how one might think about possible improvements:

- Your best guess about the average improvement will often not be a whole number (e.g. 0, 1, or 2), but instead lie between two whole numbers (e.g. 0.3 or 1.2).
- If you do not expect to change any of your answers, then the conversation will not affect your quiz score and you should answer 0.0.
- The quiz featured 14 questions. If you knew none of the answers and answered randomly, then you would still "accidentally" answer every fifth question correctly, leading to 2.8 correct answers on average.
- Suppose you expect to learn the correct answer to exactly 1 question that you answered randomly in the initial quiz and therefore had a 20 percent chance of getting

right. Then you should expect an improvement of 0.8. (Note that if there remains a chance that your new answer is wrong, then your average improvement is less than 0.8.)

- Please type in below how many more correct answers you expect to have, on average, in the revised quiz compared to the initial quiz. Please also enter the digit after the decimal point (0-9), which captures fractions of whole numbers. Please write 0.0 if you don't expect to improve at all.

Your expected improvement:

[Number between 0 and 14, up to 1 decimal point]

Now that you have given us your guess, we will flip a virtual coin. If this coin shows "heads", then you will receive the additional bonus, depending on your guess. If the coin shows "tails", you will not receive an additional bonus based on your guess. You will learn about the result of the coin flip on the next page and about whether your guess yielded a prize at the end of the experiment.

[Next Button]

Coin Flip 1

Timer

Time left to complete this page: [Countdown from 1 minute]

Your coin flip shows "heads" [tails] and you will [not] receive an additional £0.50 bonus payment if your guess is correct.

[Next Button]

Estimates 2

Timer

Time left to complete this page: [Countdown from 2 minutes and 30 seconds]

How many questions did you get correct in the initial quiz?

The quiz featured 14 questions. Suppose you knew none of the answers and chose a random response for each question, then you would still answer 2.8 questions correctly on average. If, instead you had perfect knowledge of these answers you would get all 14 correct.

How many questions do you think you answered correctly in the initial quiz? (Please provide an integer between 0 and 14).

[Integer from 0 to 14]

You will receive £0.10 if your guess is correct.

How certain are you about your expected improvement?

You previously stated that your expected improvement in the revised quiz compared to the initial quiz was [previous response]. We would now like to elicit your confidence in this prediction, by asking how certain you are that the actual improvement falls within 1.1 questions of your answer.

How likely do you think it is that the actual improvement is larger than [previous response -1.1] and smaller than [previous response $+1.1$] (please enter a number between 0-100)?

[Number from 0 to 100]

Note that stating a higher percentage indicates a greater certainty in your expected improvement.

We have two more questions for you.

To what extent do you agree with the following statements about the likely effect of the conversation with [Partner's name] on your answers in the revised quiz?

In general, my conversation partner is likely to know more than me about the topics of the quiz, so I stand to learn from my conversation partner.

[Strongly agree; Agree; Disagree; Strongly disagree]

My conversation partner is likely to know more specifically on questions that I don't know the answer to, so I stand to learn from my conversation partner.

[Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree]

[Next Button]

THERE IS JUST ONE MORE TWIST!

Timer

Time left to complete tis page: [Countdown from 2 minutes]

There is some probability (5 percent) that the computer selects you to be a decider. If the computer selects you to be a decider, then you will be able to decide whether you want to have the 8-minute conversation with [Partner's name] ([Partner's party]) or whether you want to skip the conversation altogether.

Note that the computer will only ever select you to be a decider in the revised quiz mode.

[Next Button]

DECIDER

Timer
Time left to complete this page: [Countdown from 8 minutes]

You are set to receive a £8.00 show-up bonus and to have a conversation with [Partner's name] ([Partner's party]). We now want you to consider, as an alternative, skipping the conversation altogether and, instead, spending 8 minutes on a wait screen.

Suppose throughout this section that you are in the Revised Quiz Mode. This means that the conversation with [Partner's name] ([Partner's party]) will impact your earnings from the quiz if you expect to learn from the conversation.

For your convenience, we will remind you of your answer about how much you expect to improve your quiz from the conversation: you previously told us that you expect to answer [previous response] more questions correctly in the revised quiz, each of which would increase your payment by £0.50.

Note that even someone who does not think that they will improve their earnings on the quiz, might still prefer to spend 8 minutes having a conversation with her conversation partner rather than spending 8 minutes on a wait screen. Others might prefer to spend 8 minutes on the wait screen rather than having a 8 minute conversation with their conversation partner, even if the conversation could improve their earnings on the quiz.

All things considered, if you definitely don't want to have the conversation with your partner, you will be willing to forgo at least a small amount of your show-up bonus in order to skip the conversation. Conversely, if you definitely want to have the conversion, you will require additional payment on top of your show-up bonus to skip the conversation. You might also be indifferent between having and skipping the conversation.

Which of the following applies to you? It is in your interest to answer this question as honestly as you can. As you will see in a moment, the procedure that awards payments is designed in such a way that the best thing for you to do is to honestly report your preference.

- ☐ *[I would be willing to pay at least a small part of my £8.00 show-up bonus in order to skip the conversation with [Partner's name] ([Partner's party])]*
- ☐ *[I would need to be paid at least a small additional payment on top of my £8.00 show-up bonus in order to skip the conversation with [Partner's name] ([Partner's party])]*
- ☐ *[I am indifferent between having or skipping the conversation. That is, neither am I willing to pay some of my £8.00 show-up bonus in order to skip the conversation, nor would I need to be paid an additional payment on top of my £8.00 show up bonus to skip the conversation.]*

If willing to pay to skip

You said that you would be willing to forego at least a small part of your £8.00 show-up bonus in order to skip the conversation with [Partner's name] ([Partner's party]). What is the largest amount out of the £8.00 show-up bonus that you would be willing to forego in order to skip the conversation with [Partner's name] ([Partner's party])?

Once again, it is in your interest to answer this question honestly. As you will see in a moment, the procedure that awards payments is designed in such a way that the best thing for you to do is to honestly report your preference.

Choose a number from 0 to 8, including up to one decimal point (e.g., 1.2)

[Number from 0 to 8, up to 1 decimal point]

If needs to be paid to skip

You said that you would need to be paid at least a small additional payment on top of your £8.00 show-up bonus to skip the conversation with [Partner's name] ([Partner's party]). What is the smallest additional payment on top of the £8.00 show-up bonus that you would need to receive in order to compensate you for skipping the conversation with [Partner's name] ([Partner's party])?

Once again, it is in your interest to answer this question honestly. As you will see in a moment, the procedure that awards payments is designed in such a way that the best thing for you to do is to honestly report your preference.

Choose a number from 0 to 8, including up to one decimal point (e.g., 1.2).

[Number from 0 to 8, up to 1 decimal point]

Shown after answering to questions above

Thank you for your response! We are interested in your minimum alternative bonus. Your minimum alternative bonus is defined as the alternative bonus that would make you indifferent between:

- the £8.00 show-up bonus and having the conversation with [Partner's name] ([Partner's party]).

- the alternative bonus and not having the conversation with [Partner's name] ([Partner's party]).

Good news! Based on your previous answers we already know your exact minimum alternative bonus, and we pre-filled the appropriate response for you below.

Your decision will be implemented as follows. The computer randomly picks an alternative bonus offer. If the offer is below your minimum alternative bonus, then nothing changes, i.e. you receive £8.00 and have the conversation. If the offer is above your minimum alternative bonus, then you receive the computer's offered bonus and skip the conversation.

We understand the procedure above is a bit complicated. Don't worry! Simply know that it is in your best interest to state the alternative bonus that makes you indifferent between having and not having the conversation.

- People who would rather not have the conversation state a minimum alternative bonus below £8.00, because they are willing to give up some of their show-up bonus in order to skip the conversation.
- People who want to have the conversation state a minimum alternative bonus above £8.00, because they need some additional payment on top of their show-up bonus to compensate them for skipping the conversation.
- People who are completely indifferent between having and not having the conversation would state a minimum alternative bonus of £8.00.

The choice you make below will only be implemented if you are in the Revised Quiz Mode and are selected to be a decider. Remember: in the Revised Quiz Mode having the conversation impacts your earnings from the quiz. Also note that, in most cases, you will not know whether you are in the initial or the revised quiz mode until after the experiment.

This next decision might count!

For the revised quiz mode, please select the minimum alternative bonus to skip the conversation. Based on your previous answers we already know your minimum alternative bonus and filled in a response for you. Nonetheless, you can change your answer by moving the slider if you want to.

[Slider from 0 to 16, with 0 labeled as Willing to give up some of the show-up bonus to skip conversation; 16 labeled as Need additional payment on top of show-up bonus to skip conversation; values pre-filled based on responses above]

Please tell us in two to three sentences what considerations factored into your choices about whether or not you would have the conversation?

[Open text]

Decider outcome (shown to non-decider groups)

Timer
[Countdown from 1 minute]

The computer selected you to have a conversation. Please have the conversation and complete all parts of the survey to receive your payment (between £8 and £21). And please note that not doing so seriously compromises the usefulness of this study.

Decider outcome (shown to decider groups)

If participant is a decider and conversation will take place

You were selected to be a decider. The computer implemented your choice to have the conversation.

If participant is a decider and conversation will not take place

You were selected to be a decider. The computer implemented your choice not to have the conversation.

If participant is not a decider and conversation will take place

Your partner selected to chat with you. Therefore, you have a chance to revise your answers afterwards.

If participant is not a decider and conversation will not take place

Your partner selected not to chat with you. Therefore, you cannot revise the quiz.

[Next Button]

Conversation Wait Page

Please wait actively on this Page for the other participant. If you're waiting for longer than three minutes, contact us via the chat on the bottom right and refresh this page every 30 seconds while we check what's taking your partner so long.

FACE-TO-FACE CONVERSATION

Timer

[Countdown from 8 minutes]

You can now see your conversation partner. If not, please grant permission to use your microphone and camera in your browser. Please use these 8 minutes to discuss your answers with your partner and try to figure out what the correct answers are.

Next to each statement, you can take notes on how you may revise the quiz later. These notes will be available to you when revising.

[Video frame for conversation]

Question	Your Answers	Notes
[Question and Options 1]	*[Previous Answer 1]*	*[Open text]*
[Question and Options 1]	*[Previous Answer 1]*	*[Open text]*
	⋮	
[Question and Options 14]	*[Previous Answer 14]*	*[Open text]*

REVISED QUIZ

Timer

Time left to complete this page: [Countdown from 10 minutes]

For each of the fourteen questions below, please determine the correct option. If this revised quiz is selected for payment, then you will receive £0.50 for each correct answer. Whether the revised quiz determines your payments is randomly determined by the computer and you will learn which mode you are in by the end of the experiment. With a probability of 50 percent the computer puts you in Revised Quiz mode, which makes the revised quiz count.

For each question below, please also state how confident you are in your answer, i.e. how likely it is that your answer is correct. For example, if you are very uncertain, then you might state a likelihood that your answer is correct between 20 and 30 percent. If you are very certain, then you may pick a likelihood above 90 percent.

Finally, for each question, please state the share of Republicans and the share of Democrats from a previous survey that you think answered the question correctly. You will receive £0.10 if you are within 5 percentage points of the correct share for one randomly selected question.

To review your previous answer and the notes taken during your conversation, click on the [brain symbol] button next to the corresponding fact.

You can only proceed if all answers are given and confidence levels are between 20 and 100.

Question		**Your Answer**	**Your Confidence Level**	**Correct Republicans**	**Correct Democrats**
[Question 1]	[brain symbol]	*[Options 1]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
[Question 2]	[brain symbol]	*[Options 2]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
		⋮			
[Question 14]	[brain symbol]	*[Options 14]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*

[When brain symbol is clicked, participant's answer in the Initial Quiz and notes taken during the conversation are displayed below the given question]

[Next Button]

MAKE A GUESS

Next, you will have to make a guess.

Now that you've had the conversation with [Partner's name] ([Partner's party]) and revised your quiz, please give us your best guess of how many more correct answers you expect to have in the revised quiz compared to the initial quiz.

You may win an additional £0.50 bonus for your guess. Of course, many different outcomes are possible, depending on how [Partner's name] turned out to be and how the conversation went. You maximize your chance of winning a £0.50 bonus by simply stating your best guess of how much you will improve on average. The mechanism we use to determine whether you win the bonus is a little complicated, but assures that you have a higher chance of winning £0.50 the closer your guess about the average improvement is to the true improvement.

Explanation of the mechanism in more detail (optional)

[Initially hidden, shown when clicked]

How your answer to this question maps into your chance of winning the bonus is based on a formula. This formula is designed to make sure that you maximize your chance of winning if you report your true belief of your average improvement on the quiz (therefore, you will likely need to use decimals).

The variable R is the actual improvement, comparing your correct answers on the revised quiz to the original quiz after the experiment. The variable r is what you report for your expectation of the average improvement on the question below. The winning probability for the bonus is then given by winning probability $= 100 - 100 \times (\frac{R-r}{28})^2$.

This probability is higher, the closer is your report to the actual improvement.

Please consider the following points carefully, as they speak to how one might think about possible improvements:

- Your best guess about the average improvement will often not be a whole number (e.g. 0, 1, or 2), but instead lie between two whole numbers (e.g. 0.3 or 1.2).
- If you did not change any of your answers, then the conversation did not affect your quiz score and you should answer 0.0.
- The quiz featured 14 questions. If you knew none of the answers and answered randomly, then you would still "accidentally" answer every fifth question correctly, leading to 2.8 correct answers on average.
- Suppose you learned the correct answer to exactly 1 question that you answered randomly in the initial quiz and therefore had a 20 percent chance of getting right. Then you should expect an improvement of 0.8. (Note that if there remains a chance that your new answer is wrong, then your average improvement is less than 0.8.)

Below, please type in how many more correct answers you expect to have in the revised quiz compared to the initial quiz. Please also enter the digit behind the decimal point (0-9), which captures fractions of whole numbers. Please write 0.0 if you don't expect to improve at all.

Just as a reminder, when we asked you the same question before the conversation with [Partner's name], your answer was: [previous response]

Your expected improvement:

[Number between 0 and 14, up to 1 decimal point]

Now that you have given us your guess, we will flip a virtual coin. If this coin shows "heads", then you will receive the additional bonus, depending on your guess.

[Next Button]

Coin Flip 2

Your coin flip shows "heads" [tails] and you will [not] receive an additional £0.50 bonus payment if your guess is correct.

[Next Button]

Conversation Feedback

How many questions did you get correct in the revised quiz?

The quiz featured 14 questions. Suppose you knew none of the answers and chose a random response for each question, then you would still answer 2.8 questions correctly on average. If, instead you had perfect knowledge of these answers you would get all 14 correct.

When you were asked the same question for the initial quiz, you thought that you got [previous response] correct answers.

How many questions questions do you think you answered correctly in the revised quiz? (Please provide an integer between 0 and 14).

[Integer from 0 to 14]

You will receive £0.10 if your guess is correct.

How certain are you about your expected improvement?

You previously stated that your expected improvement in the revised quiz compared to the initial quiz was [previous response]. We would now like to elicit your confidence in this prediction, by asking how certain you are that the actual improvement falls within 1.1 questions of your answer.

How likely do you think it is that the actual improvement is larger than [previous response -1.1] and smaller than [previous response $+1.1$] (please enter a number between 0-100)?

[Number from 0 to 100]

Note that stating a higher percentage indicates a greater certainty in your expected improvement.

Please answer the following questions about the conversation.

To what extent do you agree with the following statements about the conversation with [Partner's name]?

- In general, my conversation partner knew more than me about the topics of the quiz, so I learned from them.
- My conversation partner knew more specifically on questions that I did not know the answer to, so I learned from them.
- The conversation with my conversation partner got heated and confrontational.
- I did not trust my conversation partner on issues that they believed to be knowledgeable about.

- My conversation partner seemed not to trust me on issues that I believe to be knowledgeable about.
- I had a good time talking with my conversation partner.
- My conversation partner had a good time talking to me.

[Strongly agree; Agree; Disagree; Strongly disagree]

[Next Button]

Emotion Response

Now we would like to understand how you feel about the conversation you just had with your partner. Please concentrate on your feelings rather than your thoughts. Would you say that the conversation made you feel:

- Angry
- Contemptuous
- Hopeful
- Proud
- Uneasy
- Bitter
- Anxious
- Enthusiastic
- Afraid

- Disgusted

  [Very; Somewhat; Not very; Not at all]

  [Next Button]

Reasons to Converse

Thank you, you are almost done. The remainder of the study should take no longer than 5 minutes.

For each of the three questions below, you will receive £0.10 if your guess is within 10 percentage points of the correct answer. To provide more precise answers, please also enter the digit behind the decimal point (0-9), which captures fractions of whole numbers. You may enter answers between 0.0 and 14.0.

How many out of 14 questions do you think [Partner's name] answered correctly in the initial quiz, before the conversation?

[Number between 0 and 14, up to 1 decimal point]

How many out of 14 answers to the quiz do you think [Partner's name] changed after talking to you?

[Number between 0 and 14, up to 1 decimal point]

How many out of 14 questions do you think [Partner's name] answered correctly in the revised quiz, after the conversation? (Keep in mind that changing answers doesn't necessarily imply having the correct answer)

[Number between 0 and 14, up to 1 decimal point]

Did the following considerations affect your willingness to have the conversation with [Partner's name] today. (Please answer these questions disregarding how the conversation actually went and focus instead on your considerations prior to joining the conversation.)

- Improving your own score on the quiz.
- Having a good time with a fellow study participant.
- Helping my conversation partner do better on their quiz.
- Helping spread the word to my conversation partner about the right position on policy issues.
- Learning about the perspective of another person.
- Worrying about the conversation being uncomfortable.
- Worrying about the conversation being annoying.
- Worrying about the technology not working.

[Not at all; To a small extent; To some extent; To a large extent]

[Next Button]

Bonus Quiz

Timer

Time left to complete this page: [Countdown from 3 minutes]

You can take up to 3 minutes to complete this quiz.

For each of the three questions below, please determine the correct option. You will receive £0.50 for each correct answer.

For each question below, please also state how confident you are in your answer, i.e. how likely it is that your answer is correct. For example, if you are very uncertain, then you might state a likelihood that your answer is correct between 20 and 30 percent. If you are very certain, then you may pick a likelihood above 90 percent.

Finally, for each question, please state the share of Republicans and the share of Democrats from a previous survey that you think answered the question correctly.

> You can only proceed if all answers are given and confidence levels are between 20 and 100.

Question	**Your Answer**	**Your Confidence Level**	**Correct Republicans**	**Correct Democrats**
[B Question 1]	*[B Options 1]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
[B Question 2]	*[B Options 2]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*
[B Question 3]	*[B Options 3]*	*[Number 20-100]*	*[Number 0-100]*	*[Number 0-100]*

[Next Button]

Risk Aversion

Note: the returns from the task below will actually be added to the final payoff for 1 in 10 participants.

We are now giving you an additional 200 pence (£2.00), on top of the earnings you have accumulated so far. You decide how many of these pence (between 0 and 200, inclusive) to invest. Those pence not invested are yours to keep.

The investment has a 50 percent chance of success. To determine if your investment is a success, a fair coin is tossed:

- If the coin comes up heads, your investment pays 2.5 times the amount you invested.
- If the coin comes up tails, you lose the amount invested.

Examples:

1. You invest 200 pence (all of your £2.00).
   - If the coin comes up heads, you will earn 500 pence (£5.00) from your investment. Thus, you bring home £5.00 in total.
   - If it comes up tails, you will lose your £2.00 investment. Thus, you bring home £0.00.
2. You invest 100 pence (£1.00).
   - If the coin comes up heads, you will earn 250 pence (£2.50) from your investment. Combined with the £1.00 you didn't invest, you bring home £3.50 in total.
   - If it comes up tails, you will lose your £1.00 investment but keep the other £1.00. Thus, you bring home £1.00.
3. You decide not to invest any pence. Regardless of the coin toss, you will keep your £2.00, so you bring home £2.00.

At the end of the experiment, your returns from this task will be determined based on your investment decision and the coin toss result. There is a 1 in 10 chance that this amount will be added to your final payoff.

Please indicate how many pence you want to invest:

[Integer from 0 to 200]

Enter a value between 0 and 200 pence (0 to £2.00).

[Next Button]

Affective Polarization Follow-up Survey

Thank you, we are almost done! We just have a few more general questions.

We would like to get your feelings toward specific groups of people. We will show the name of a group and we would like you to rate the group using something we call the feeling thermometer.

Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 degrees and 50 degrees mean that you do not feel favorable toward the group and that you do not care too much for that group. You would rate the group at the 50 degree mark if you do not feel particularly warm or cold toward the group.

How would you rate **Democrats**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

How would you rate **Republicans**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

How well do the following statements describe you:

- I find it difficult to see things from Democrats' point of view
- I find it difficult to see things from Republicans' point of view
- I think it is important to consider the perspective of Democrats
- I think it is important to consider the perspective of Republicans

[Not at all; Slightly; Moderately; Very; Extremely]

In your judgment, the **Democratic** party has

[A lot of good ideas; Some good ideas; A few good ideas; Almost no good ideas]

In your judgment, the **Republican** party has

[A lot of good ideas; Some good ideas; A few good ideas; Almost no good ideas]

How would you feel if you had a daughter or son who married a **[Member of opposite party]**?

[Very upset; Somewhat upset; Not upset at all]

[Next Button]

News Consumption

In the past 24 hours have you... (check all that apply)

☐ Used social media (such as Facebook or Youtube)

☐ Watched TV news

☐ Read a newspaper in print or online

☐ Listened to a radio news program or talk radio

☐ None of these

In the past 24 hours, did you do any of the following on social media (such as Facebook, Youtube or Twitter)? (check all that apply)

☐ Posted a story, photo, video or link about politics

☐ Posted a comment about politics

☐ Read a story or watched a video about politics

☐ Followed a political event

☐ Forwarded a story, photo, video or link about politics to friends

☐ None of the above

On an average day of the last 4 weeks, how many minutes did you spend talking about local, political, or social issues with someone **who presumably votes Democrat**?

[Number]

On an average day of the last 4 weeks, how many minutes did you spend talking about local, political, or social issues with someone **who presumably votes Republican**?

[Number]

On a scale from 1 to 10, where 1 is "not at all important" and 10 means "absolutely important", how important is it for you to live in a country that is governed democratically?

[Integer from 1 to 10]

On a scale from 1 to 10, where 1 is "not at all hopeful" and 10 means "absolutely hopeful", how hopeful are you about your country's democratic institutions to master the challenges of the coming decades?

[Integer from 1 to 10]

[Next Button]

Personality Traits

Do the following statements apply to you?

1. I am quite effective at talking people into things.
2. I am quite good at convincing others to see things my way.
3. I am very good at influencing people.
4. I do not like to be the center of attention on social occasions.
5. I do not like to organize other people's activities.
6. I don't enjoy trying to convince people of something.
7. I enjoy being in the spotlight.
8. I perform for an audience whenever I can.

9. I usually do not like to be a ”follower.”
10. In most social situations I like to have someone else take the lead.
11. In social situations I usually allow others to dominate the conversation.
12. People find me forceful.
13. When I work with others I like to take charge.
14. When it is time to make decisions, others usually turn to me.

[True; False]

[Next Button]

Final Survey

This is the last set of questions.

What year were you born?

[Year]

Which state do you live in?

[Dropdown list of possible states]

What is your educational attainment?

[No high school graduation; High school graduate; Some college, but no degree; Associate's degree; Bachelor's degree; Graduate or professional degree]

What's your approximate annual household income?

[Below 10,000; 10,000—14,999; 15,000—24,999; 25,000—34,999; 35,000—49,999; 50,000—74,999; 75,000—99,999; 100,000—149,999; 150,000—199,999; 200,000 or more]

How many people live in your household?

[Integer]

How would you describe your current place of residence?

[Mega city: population more than one million; Big city: population between 200,000 and one million; City: population between 50,000 and 200,000; Small city near metropolitan area: population between 20,000 and 50,000, close to metro area; Small city not near metropolitan area: population between 20,000 and 50,000, not close to metro area; Town near metropolitan area: population between 3,000 and 20,000, close to metro area; Town not near metropolitan area: population between 3,000 and 20,000, not close to metro area; Village near metropolitan area: population between 500 and 3,000, close to metro area; Village not near metropolitan area: population between 500 and 3,000, not close to metro area; Small village: population less than 500]

Are you of hispanic, Latino, or Spanish origin?

[Yes; No]

How would you define your race category? You may choose one or more races. For this survey, Hispanic origin is not a race.

☐ White

☐ Black or African American

☐ American Indian or Alaska Native

☐ Asian

☐ Native Hawaiian or Other Pacific Islander

What is your current employment status?

[Full time employee; Part-time employee; Self-employed or small business owner; Unemployed and looking for work; Student; Not in labor force (for example: retired or full-time parent)]

Who did you vote for in the 2020 presidential election?

[Donald Trump; Joe Bidenl; Other; Did not vote]

We hear a lot of talk these days about liberals and conservatives. Here is a seven-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale, or haven't you thought much about this?

[Very Liberal; Liberal; Somewhat Liberal; Neutral; Somewhat Conservative; Conservative; Very Conservative; I have not thought much about it]

[Next Button]

Payoff

Your payoff for this experiment is now being calculated and you will be getting this as a bonus on Test within 7 days.

You can now close this tab.

J. Instructions: Follow-up Survey

Consent

Welcome! In this study we will ask for your personal assessments.

Data is being collected confidentially for research purposes. At any point, even after the survey is submitted, you may contact Luca Braghieri (luca.braghieri@unibocconi.it) to voice your concerns and or withdraw your consent.

Please make your choice below:

[Yes, I agree to participate in this study; No, I do not agree to participate]

Feeling thermometer

We'd like to get your feelings toward specific groups of people. We'll show the name of a group and we'd like you to rate that group using something we call the feeling thermometer.

Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 degrees and 50 degrees mean that you do not feel favorable toward the group and that you don't care too much for that group. You would rate the group at the 50 degree mark if you don't feel particularly warm or cold toward the group.

How would you rate **Democrats**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

How would you rate **Republicans**?

[Slider from 0 to 100, where 0 is labeled as "Very cold or unfavorable feeling"; 50 as "No feeling at all"; 100 as "Very warm or favorable feeling"]

Affective polarization II

How well do the following statements describe you:

- I find it difficult to see things from Democrats' point of view
- I find it difficult to see things from Republicans' point of view
- I think it is important to consider the perspective of Democrats
- I think it is important to consider the perspective of Republicans

[Not at all; Slightly; Moderately; Very; Extremely]

In your judgment, the **Democratic** party has

[A lot of good ideas; Some good ideas; A few good ideas; Almost no good ideas]

In your judgment, the **Republican** party has

[A lot of good ideas; Some good ideas; A few good ideas; Almost no good ideas]

How would you feel if you had a daughter or son who married a **[Member of opposite party]**?

[Very upset; Somewhat upset; Not upset at all]

Willingness to interact

You previously participated in an experiment we conducted that had you answer a quiz on political issues and then discuss your answers with another participant via video call.

Researchers like us are sometimes asked to run extensions of their study. With some likelihood, we will run a new study in which we may ask you to answer a quiz like the one in the previous experiment on issues that are of central importance to the upcoming elections, and to meet a **[Member of opposite party]** on a video call to discuss the quiz and potentially improve your answers.

Would you rather receive 1 US$ or have this conversation? (Note that this choice might really be implemented, if we conduct a follow-up study and you agree to participate).

[I would rather have the conversation and forfeit 1 US$; I would rather receive the 1 US$ and skip the conversation]

To what extent do you think that the following considerations affect your willingness to have the conversation with this person?

- I think that I may improve my factual knowledge from this conversation.
- I worry that it will be an annoying conversation.

[Not at all; To a small extent; To some extent; To a large extent]

Quiz

We are now going to ask you the same questions we asked you in our previous study. For each correct answer, you will receive a 5 cents bonus.

For this task, it is important that you do not do any research or use AI models (like ChatGPT). We are only interested in what you currently know.

Question	**Answer**
[Question 1]	*[Options 1]*
[Question 2]	*[Options 2]*
⋮	
[Question 14]	*[Options 14]*

Exit

Thank you for taking part in this survey. We will process your base payment as soon as possible and your bonus payment (for the quiz) within 10 days.

Please move to the next page to complete your submission.

K. Pre-registrations files

Talking across the aisle: an experiment on political discourse (#155100)

Author(s)
Luca Braghieri (Bocconi University) - luca.braghieri@unibocconi.it
Peter Schwardmann (Carnegie Mellon University) - schwardmann@cmu.edu
Egon Tripodi (Hertie School) - tripodi@hertie-school.org

Pre-registered on: 12/13/2023 02:09 AM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Our experiment investigates beliefs about, preferences over, and experiences of cross-partisan and co-partisan conversations involving policy-relevant factual knowledge between Republicans and Democrats in the US.

Our first three main hypotheses are that, despite Democrats and Republicans being on-average knowledgeable about different issues: i) participants expect to learn more from a co-partisan conversation than from a cross-partisan conversation, ii) the conversation does little to correct this expectation, iii) participants indeed learn more from co-partisan than from counter-partisan conversations.

Our fourth main hypothesis is that participants prefer having co-partisan conversations over cross- partisan conversations, net of the monetary benefits of learning.

Additional analyses of interest will focus on the determinants of how much is actually learned due to the conversation and the effect of the conversation on affective and factual polarization.

3) Describe the key dependent variable(s) specifying how they will be measured.
We measure whether Democrats and Republicans are on average knowledgeable about different political issues (distributed knowledge) by administering a political knowledge quiz and comparing, across different questions, the fraction of Democrats and Republicans who answer correctly.

We measure perceived learning before and after the conversation using an incentivized elicitation of the average improvement in the number of correct answers moving from an initial quiz, before the conversation, to a revised quiz, after the conversation. We also measure the actual improvement on the quiz as a result of the conversation.

We measure the preference for having the conversation, net of the monetary benefits of learning, as the willingness to pay for having the conversation.

4) How many and which conditions will participants be assigned to?
Participants are randomly assigned to a conversation partner. We will compare participants in cross-partisan with participants in co-partisan partnerships.

At a finer level of aggregation, we will compare participants randomly assigned to Democrat-Democrat pairs, Democrat-Republican pairs, and Republican-Republican pairs.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will test the first hypothesis by comparing expected learning for co-partisan and cross-partisan interactions before the conversation. We will employ a standard two-sided t-test for the comparison.

We will test the second hypothesis by comparing expected learning before the interaction with expected learning after the interaction as a function of whether the conversation was with a co-partisan or with a counter-partisan. We will employ standard two-sided t-tests for the comparison in a difference-in-differences framework. The single difference after the conversation will also be relevant.

We will test the third hypothesis by comparing a participant's improvement on the quiz based on the interaction as a function of whether the conversation was with a co-partisan or with a counter-partisan. We will employ standard two-sided t-tests for the comparison.

Lastly we will test the fourth hypothesis by comparing the willingness to pay for co-partisan interactions and the willingness to pay for counter-partisan interactions. We will employ standard two-sided t-tests for the comparison.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
The experiment will be run on Prolific. We will only invite participants that self-identify as either Republican or Democrat, in equal proportions. We will exclude participants who are unable to pass a check that assures that they are able to be visible and audible during a video chat.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We aim to recruit 1000 participants, which will be randomly assigned to Democrat-Democrat pairs (25%), Democrat-Republican pairs (50%), and Republican-Republican pairs (25%).

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
Additional analyses will explore how learning is predicted by features of the conversation such as its heatedness, participants' attitudes toward and beliefs about cross-partisans, and communication patterns identified from scripts of the conversations. We will also construct a variable that captures the potential for learning from the conversation based on initial answers and participants' confidence in them to serve as a benchmark.

In a separate investigation based on the same dataset, we will study the effect of cross-partisan conversations (relative to co-partisan conversations and the prior baseline) on affective and factual polarization. Our main hypotheses, tested using two-sided t-tests, for this separate investigation is that our measures of post-conversation factual and affective polarization are smaller in cross- than in co-partisan couples. We will also test how this depolarizing effect of cross-partisan contact is mediated by participants' eagerness to engage in these conversations, as measured by our willingness to pay elicitation.

Persistence in affective polarization reduction (#173633)

Author(s)
Egon Tripodi (Hertie School) - tripodi@hertie-school.org
Luca Braghieri (Bocconi University) - luca.braghieri@unibocconi.it
Peter Schwardmann (Carnegie Mellon University) - schwardmann@cmu.edu

Pre-registered on: 05/04/2024 02:44 AM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Our experiment investigates whether affective polarization is meaningfully decreased by cross-partisan contact several weeks after the contact.

3) Describe the key dependent variable(s) specifying how they will be measured.
Our primary dependent variable is derived from the answers to feelings thermometer questions, commonly used in political science research. Participants are asked how warm they are feeling toward members of their own and the other party. Each participants' partisan gap is then calculated by subtracting their feelings towards members of the other party from their feelings towards their party.

As secondary dependent variables we will also create an index of affective polarization that is based on four measures of polarization based on: feeling thermometer (own vs other party), difficulty in seeing the perspective of the own vs other side, perceived importance to consider the perspective of own vs other other side, and quality of own vs other party's ideas.

For another secondary analysis, we will also measure willingness to interact with a member of the other party in a potential future experiment.

4) How many and which conditions will participants be assigned to?
We don't introduce any experimental variation within this survey, but we instead use the experimental variation from an experiment run several weeks prior. That experiment randomized whether study participants have a conversation with co- vs cross- partisans.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will test the main hypothesis that treatment assignment in the previous experiment still has a persistent effect on affective polarization. In OLS regressions, we will regress our main dependent variables on a dummy that takes a value of 1 if the participant had a cross-partisan interaction. We will cluster standard errors at the conversation pair level. We will control for baseline characteristics.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
The experiment will be run on Prolific and Cloudresearch Connect. We will only invite the 993 participants in the original study sample.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
Ideally we will be able to re-sample at least 60 percent of participants from the original intervention, but the exact number of participants that we will be able to recruit is hard to anticipate.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
In OLS regressions, we will regress our previously specified secondary dependent variables on a dummy that takes a value of 1 if the participant had a cross-partisan interaction. We will cluster standard errors at the conversation pair level. We will control for baseline characteristics.

Appendix References

Abu-Rayya, Hisham M., "Majority members' endorsement of the acculturation integrationist orientation improves their outgroup attitudes toward ethnic minority members: An electronic-contact experiment," *Computers in Human Behavior*, October 2017, *75*, 660–666.

Adachi, Paul J. C., Gordon Hodson, Teena Willoughby, and Sarah Zanette, "Brothers and sisters in arms: Intergroup cooperation in a violent shooter game can reduce intergroup bias," *Psychology of Violence*, 2015, *5* (4), 455–462.

Albuja, Analía F., Sarah E. Gaither, Diana T. Sanchez, and Jaelyn Nixon, "Testing intergroup contact theory through a natural experiment of randomized college roommate assignments in the United States," *Journal of Personality and Social Psychology*, 2024, *127* (2), 277–290.

Alimo, Craig John, "From Dialogue to Action: The Impact of Cross-Race Intergroup Dialogue on the Development of White College Students as Racial Allies," *Equity & Excellence in Education*, January 2012, *45* (1), 36–59.

Andrews, Isaiah and Maximilian Kasy, "Identification of and Correction for Publication Bias," *American Economic Review*, August 2019, *109* (8), 2766–2794.

Asimovic, Nejla, Ruth K. Ditlmann, and Cyrus Samii, "Estimating the effect of intergroup contact over years: evidence from a youth program in Israel," *Political Science Research and Methods*, July 2024, *12* (3), 475–493.

Baron, Hannah, Robert Blair, Donghyun Danny Choi, Laura Gamboa, Jessica Gottlieb, Amanda Lea Robinson, Steven Rosenzweig, Megan Turnbull, and Emily A. West, "Couples Therapy for a Divided America: Assessing the Effects of Reciprocal Group Reflection on Partisan Polarization," May 2021.

Benatov, Joy, Rony Berger, and Carmit T. Tadmor, "Gaming for peace: Virtual contact through cooperative video gaming increases children's intergroup tolerance in the context of the Israeli–Palestinian conflict," *Journal of Experimental Social Psychology*, January 2021, *92*, 104065.

Berger, Rony, Joy Benatov, Hisham Abu-Raiya, and Carmit T. Tadmor, "Reducing prejudice and promoting positive intergroup attitudes among elementary-school children in the context of the Israeli–Palestinian conflict," *Journal of School Psychology*, 2016, *57*, 53–72.

Bilewicz, Michal, "History as an obstacle: Impact of temporal-based social categorizations on Polish-Jewish intergroup contact," *Group Processes & Intergroup Relations*, 2007, *10* (4), 551–563.

_ **and Manana Jaworska**, "Reconciliation through the Righteous: The Narratives of Heroic Helpers as a Fulfillment of Emotional Needs in PolishJewish Intergroup Contact," *Journal of Social Issues*, 2013, *69* (1), 162–179.

Boag, Elle Mae and David Wilson, "Inside experience: engagement empathy and prejudice towards prisoners," *Journal of Criminal Psychology*, January 2014, *4* (1), 33–43.

Boccanfuso, Emery, Fiona A. White, and Rachel D. Maunder, "Reducing transgender stigma via an e-contact intervention," *Sex Roles: A Journal of Research*, 2021, *84* (5-6), 326–336.

Boersma, Paul and David Weenink, "Praat: doing phonetics by computer [Computer program]," Version 6.1.38, retrieved 2 January 2021 http://www.praat.org/ 2021.

Broockman, David and Joshua Kalla, "Durably reducing transphobia: A field experiment on door-to-door canvassing," *Science*, April 2016, *352* (6282), 220–224.

Chang, Han Il and Leonid Peisakhin, "Building Cooperation among Groups in Conflict: An Experiment on Intersectarian Cooperation in Lebanon," *American Journal of Political Science*, 2019, *63* (1), 146–162.

Combs, Aidan, Graham Tierney, Brian Guay, Friedolin Merhout, Christopher A. Bail, D. Sunshine Hillygus, and Alexander Volfovsky, "Reducing political polarization in the United States with a mobile chat platform," *Nature Human Behaviour*, September 2023, *7* (9), 1454–1461.

Cook, Anna, Jane Ogden, and Naomi Winstone, "The impact of a school-based musical contact intervention on prosocial attitudes, emotions and behaviours: A pilot trial with autistic and neurotypical children," *Autism*, May 2019, *23* (4), 933–942.

Corno, Lucia, Eliana La Ferrara, and Justine Burns, "Interaction, Stereotypes, and Performance: Evidence from South Africa," *American Economic Review*, December 2022, *112* (12), 3848–3875.

Cramwinckel, Florien M., Daan T. Scheepers, Tom F. Wilderjans, and Robert-Jan B. de Rooij, "Assessing the Effects of a Real-Life Contact Intervention on Prejudice Toward LGBT People," *Archives of Sexual Behavior*, 2021, *50* (7), 3035–3051.

DellaVigna, Stefano and Elizabeth Linos, "RCTs to scale: Comprehensive evidence from two nudge units," *Econometrica*, 2022, *90* (1), 81–116.

Dessel, Adrienne B., "Effects of Intergroup Dialogue: Public School Teachers and Sexual Orientation Prejudice," *Small Group Research*, October 2010, *41* (5), 556–592.

Elwert, Felix, Tamás Keller, and Andreas Kotsadam, "Rearranging the desk chairs: A large randomized field experiment on the effects of close contact on interethnic relations," *American Journal of Sociology*, 2023, *128* (6), 1809–1840.

Ermer, Ashley E., Katie York, and Katharine Mauro, "Addressing ageism using intergenerational performing arts interventions," *Gerontology & Geriatrics Education*, 2021, *42* (3), 308–315.

Gabrielli, Sara, Maria Gaetana Catalano, Fridanna Maricchiolo, Daniele Paolini, and Paola Perucchini, "Reducing implicit prejudice towards migrants in fifth grade pupils: efficacy of a multi-faceted school-based program," *Social Psychology of Education*, June 2022, *25* (2), 425–440.

Gaither, Sarah E. and Samuel R. Sommers, "Living with an other-race roommate shapes Whites' behavior in subsequent diverse settings," *Journal of Experimental Social Psychology*, 2013, *49* (2), 272–276.

Goette, Lorenz and Egon Tripodi, "The limits of social recognition: Experimental evidence from blood donors," *Journal of Public Economics*, 2024, *231*, 105069.

Graham, Heather E., Mark C. Frame, and Jared B. Kenworthy, "The moderating effect of prior attitudes on intergroup face-to-face contact," *Journal of Applied Social Psychology*, 2014, *44* (8), 547–556.

Gu, Jun, Annika Mueller, Ingrid Nielsen, Jason Shachat, and Russell Smyth, "Improving Intergroup Relations through Actual and Imagined Contact: Field Experiments with Malawian Shopkeepers and Chinese Migrants," *Economic Development and Cultural Change*, October 2019, *68* (1), 273–303.

_ , Ingrid Nielsen, Jason Shachat, Russell Smyth, and Yujia Peng, "An experimental study of the effect of intergroup contact on attitudes in urban China," *Urban Studies*, 2016, *53* (14), 2991–3006.

Guyer, Joshua J, Leandre R Fabrigar, and Thomas I Vaughan-Johnston, "Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion," *Personality and Social Psychology Bulletin*, 2019, *45* (3), 389–405.

Harper, Eric, Somshubra Majumdar, Oleksii Kuchaiev, Li Jason, Yang Zhang, Evelina Bakhturina, Vahid Noroozi, Sandeep Subramanian, Nithin Koluguri, Jocelyn Huang, Fei Jia, Jagadeesh Balam, Xuesong Yang, Micha Livne, Yi Dong, Sean Naren, and Boris Ginsburg, "NeMo: a toolkit for Conversational AI and Large Language Models," 2024.

Hatoum, Amaani H. and Fiona A. White, "Advancing E-contact to Reduce Intergroup Anxiety and Increase Positive Attitudes Towards Individuals Who Identify as Bisexual," *Journal of Sex Research*, September 2022, *59* (7), 872–885.

Hobolt, Sara B., Katharina Lawall, and James Tilley, "The Polarizing Effect of Partisan Echo Chambers," *American Political Science Review*, August 2024, *118* (3), 1464–1479.

Holoien, Deborah Son, Hilary B. Bergsieker, J. Nicole Shelton, and Jan Marie Alegre, "Do you really understand? Achieving accuracy in interracial relationships," *Journal of Personality and Social Psychology*, January 2015, *108* (1), 76–92.

Influs, Moran, Maayan Pratt, Shafiq Masalha, Orna Zagoory-Sharon, and Ruth Feldman, "A social neuroscience approach to conflict resolution: Dialogue intervention to Israeli and Palestinian youth impacts oxytocin and empathy," *Social Neuroscience*, August 2019, *14* (4), 378–389.

—, **Shafiq Masalha, Orna Zagoory-Shaon, and Ruth Feldman**, "Dialogue intervention to youth amidst intractable conflict attenuates stress response to outgroup," *Hormones and Behavior*, April 2019, *110*, 68–76.

Jadoul, Yannick, Bill Thompson, and Bart de Boer, "Introducing Parselmouth: A Python interface to Praat," *Journal of Phonetics*, 2018, *71*, 1–15.

Kende, Anna, Linda Tropp, and Nóra Anna Lantos, "Testing a contact intervention based on intergroup friendship between Roma and non-Roma Hungarians: Reducing bias through institutional support in a non-supportive societal context," *Journal of Applied Social Psychology*, 2017, *47* (1), 47–55.

Koball, Afton M. and Robert A. Carels, "Intergroup contact and weight bias reduction," *Translational Issues in Psychological Science*, 2015, *1* (3), 298–306.

Levendusky, Matthew S. and Dominik A. Stecula, "We Need to Talk: How Cross-Party Dialogue Reduces Affective Polarization," *Elements in Experimental Political Science*, November 2021.

Levy, Jonathan, Moran Influs, Shafiq Masalha, Abraham Goldstein, and Ruth Feldman, "Dialogue intervention for youth amidst intractable conflict attenuates neural prejudice response and promotes adults' peacemaking," *PNAS Nexus*, November 2022, *1* (5), pgac236.

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, June 2021, *111* (6), 1807–1844.

Lytle, Ashley and Sheri R. Levy, "Reducing heterosexuals' prejudice toward gay men and lesbian women via an induced cross-orientation friendship," *Psychology of Sexual Orientation and Gender Diversity*, 2015, *2* (4), 447–455.

MacKuen, Michael, Jennifer Wolak, Luke Keele, and George E. Marcus, "Civic Engagements: Resolute Partisanship or Reflective Deliberation," *American Journal of Political Science*, 2010, *54* (2).

Maunder, Rachel D., Fiona A. White, and Stefano Verrelli, "Modern avenues for intergroup contact: Using E-contact and intergroup emotions to reduce stereotyping and social distancing against people with schizophrenia," *Group Processes & Intergroup Relations*, October 2019, *22* (7), 947–963.

Mousa, Salma, "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq," *Science*, August 2020, *369* (6505), 866–870.

Nägele, Sophie, Katja Corcoran, and Ulrich Klocke, "It is worth taking a closer look: A field experiment on intergroup contact between Austrian pupils and refugees," *Group Processes & Intergroup Relations*, June 2024, *27* (4), 882–902.

Ozaydin, Tuba, Deniz Kocoglu Tanyer, and Belgin Akin, "Promoting the attitudes of nursing students towards refugees via interventions based on the contact hypothesis: A randomized controlled trial," *International Journal of Intercultural Relations*, September 2021, *84*, 191–199.

Paluck, Elizabeth Levy, Roni Porat, Chelsey S Clark, and Donald P Green, "Prejudice reduction: Progress and challenges," *Annual review of psychology*, 2021, *72*, 533–560.

Petrík, Juraj and Miroslav Popper, "Contact Based School Intervention Program: Enhancing Cooperation Intention and Reducing Prejudice toward Roma," *Studia Psychologica*, October 2020, *62* (3), 232–245.

Rossiter, Erin L., "The Similar and Distinct Effects of Political and Non-Political Conversation on Affective Polarization," January 2023.

_ **and Taylor N. Carlson**, "Cross-Partisan Conversation Reduced Affective Polarization for Republicans and Democrats Even after the Contentious 2020 Election," *The Journal of Politics*, October 2024, *86* (4), 1608–1612.

Rw, Romy and Nick Joyce, "Effects of Intergroup Communication on Intergroup Anxiety and Prejudice Through Single Sessions of Peer Counseling in Online Settings," *International Journal of Communication*, December 2023, *18* (0), 24.

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

Scacco, Alexandra and Shana S. Warren, "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria," *American Political Science Review*, August 2018, *112* (3), 654–677.

Shmyrev, Nickolay V. et al., "Vosk Speech Recognition Toolkit: Offline speech recognition API for Android, iOS, Raspberry Pi and servers with Python, Java, C# and Node," 2024.

Smith, Vicki L and Herbert H Clark, "On the course of answering questions," *Journal of memory and language*, 1993, *32* (1), 25–38.

Stiff, Chris and Tom Bowen, "Two-player game: Playing casual video games with outgroup members reduces levels of prejudice toward that outgroup," *International Journal of Human-Computer Interaction*, 2016, *32* (12), 912–920.

Tal-Or, Nurit and Yariv Tsfati, "When Arabs and Jews Watch TV Together: The Joint Effect of the Content and Context of Communication on Reducing Prejudice," *Journal of Communication*, 2016, *66* (4), 646–668.

Tausch, Nicole, Michèle Denise Birtel, Paulina Górska, Sidney Bode, and Carolina Rocha, "The Causal Effect of an Intergroup Contact Intervention on Affective Polarization around Brexit: A Randomized Controlled Trial," June 2023.

Voelkel, Jan G., Dongning Ren, and Mark J. Brandt, "Inclusion reduces political prejudice," *Journal of Experimental Social Psychology*, 2021, *95.*

Walch, Susan E., Kimberly A. Sinkkanen, Elisabeth M. Swain, Jacquelyn Francisco, Cassi A. Breaux, and Marie D. Sjoberg, "Using intergroup contact theory to reduce stigma against transgender individuals: Impact of a transgender speaker panel presentation," *Journal of Applied Social Psychology*, 2012, *42* (10), 2583–2605.

Welker, Keith M., Richard B. Slatcher, Lynzey Baker, and Arthur Aron, "Creating positive out-group attitudes through intergroup couple friendships and implications for compassionate love," *Journal of Social and Personal Relationships*, 2014, *31* (5), 706–725.

White, Fiona A., Hisham M. Abu-Rayya, and Chela Weitzel, "Achieving twelve-months of intergroup bias reduction: The dual identity-electronic contact (DIEC) experiment," *International Journal of Intercultural Relations*, 2014, *38*, 158–163.

Yablon, Yaacov B., "Are we preaching to the converted? The role of motivation in understanding the contribution of intergroup encounters," *Journal of Peace Education*, December 2012, *9* (3), 249–263.

Yücel, Emine, Hatice Ekici, and Ayşe Betül Çelik, ""We are One Team": An evaluation of an intervention for improving relations between Turkish local and refugee adolescents," *International Journal of Intercultural Relations*, September 2023, *96*, 101865.