nature human behaviour

Article

https://doi.org/10.1038/s41562-024-02043-y

Interacting as equals reduces partisan polarization in Mexico

Received: 20 December 2022

Accepted: 2 October 2024

Published online: 11 November 2024

Check for updates

Kenneth F. Greene[1], Erin L. Rossiter[2], Enrique Seira[3] & Alberto Simpser[4]

In many contemporary democracies, political polarization increasingly involves deep-seated intolerance of opposing partisans. The decades-old contact hypothesis suggests that cross-partisan interactions might reduce intolerance if individuals interact with equal social status. Here we test this idea by implementing collaborative contact between 1,227 pairs of citizens (2,454 individuals) with opposing partisan sympathies in Mexico, using the online medium to credibly randomize participants' relative social status within the interaction. Interacting under both equal and unequal status enhanced tolerant behaviour immediately after contact; however, 3 weeks later, only the salutary effects of equal contact endured. These results demonstrate that a simple, scalable intervention that puts people on equal footing can reduce partisan polarization and make online contact into a prosocial force.

In recent years, many countries have experienced partisan polarization severe enough to undermine trust in institutions and threaten the stability of democracy[1–4]. In such environments, political divisions can align with pre-existing social cleavages, including social status. Partisan political divides that extend into society are often called 'affective polarization', a particularly insidious form of animus that decreases cross-partisan interaction and exacerbates mutual intolerance[2,5]. Could creating opportunities for people from opposing political camps to interact under conditions of equal status increase tolerance?

Social status lies at the centre of the well-known contact hypothesis. Allport[6] argued that cross-group collaboration that endows participants with equal status can increase intergroup tolerance by enhancing perceived commonality and interpersonal closeness (ch. 30). In Allport's words, "whatever makes for equal status relationships and for more intimate acquaintance is likely to make for increased tolerance" (p. 498). By contrast, contact under unequal status is less likely to bring dissonant groups closer and lacks the beneficial effects on intergroup tolerance[6–8]. Yet the hypothesized role of status equality within the contact interaction lacks experimental validation, and observational research has yielded mixed findings[9–11].

We test Allport's (1954) equal-status hypothesis by inducing contact between citizens with opposing partisan sympathies while experimentally varying their social status in the contact situation. Paired participants collaborated online on non-political tasks for 10 min. The first task asked participants to decide whether fellow citizens in general value friendship or professional success more highly; the second task consisted of trivia questions about popular culture. Paired participants were provided with a text-chat window and encouraged to communicate with each other while completing these tasks. For our main analysis, we manipulated the participants' relative status within the interaction[12]. In the equal-status condition, participants were informed that their respective answers to the tasks counted equally towards pair-level rewards. In the unequal-status condition, one participant was designated Leader and the other Follower, with only the Leader's answers determining pair-level rewards.

We tested for both immediate and long-term effects of contact on tolerant behaviours. Consistent with some previous nonexperimental research, our experiment shows that, immediately after treatment, all forms of intergroup contact enhanced tolerant behaviour towards opposing partisans[9]. However, 3 weeks after treatment, only the experience of intergroup contact under status equality continued to enhance tolerant behaviour towards opposing partisans; the salutary effects of interacting under status inequality had disappeared (Supplementary Table 13). We focus on these longer-term effects because they are less frequently measured in experimental research and are more relevant for policy.

[1]Department of Government, University of Texas at Austin, Austin, TX, USA. [2]Department of Political Science, University of Notre Dame, Notre Dame, IN, USA. [3]Department of Economics, Michigan State University, East Lansing, MI, USA. [4]Department of Political Science, ITAM, Mexico City, Mexico.
e-mail: alberto.simpser@itam.mx

Three weeks after contact, we find that participants assigned to the equal-status condition were willing to share 24% more of their own cash points with an anonymous study participant of opposing partisanship in a dictator game, compared with those in the no-contact control group ($\beta = 24.11$; $P = 0.010$; 95% confidence interval (CI), 5.78, 42.45). Participants in the equal-status condition were also 5 percentage points (pp) more willing to accept an invitation to a future 30 min meeting to discuss the country's problems with a group of people that they were told would include opposing partisans ($\beta = 0.05$; $P = 0.009$; 95% CI, 0.01, 0.08). A standardized index of tolerant behaviour that combines sharing and willingness to dialogue was 0.17 standard deviations greater under equal-status contact compared with no contact ($\beta = 0.17$; $P < 0.001$; 95% CI, 0.08, 0.26) and 0.13 standard deviations greater under equal- compared with unequal-status contact (unequal contact $\beta = 0.04$; $P = 0.47$; 95% CI, −0.07, 0.15; equal minus unequal status difference $P = 0.009$; 95% CI, 0.03, 0.23). In the unequal-status condition, tolerant behaviour did not differ significantly for either Leaders or Followers compared with no contact (Leaders $\beta = 0.05$; $P = 0.46$; 95% CI, −0.09, 0.19; Followers $\beta = 0.03$; $P = 0.71$; 95% CI, −0.11, 0.17). Analysis of chat content suggests that the overall quality of interaction was lower between participants assigned to the inequality condition in relation to the equal-status one.

Moving beyond the contact hypothesis, we additionally test whether knowledge of real-world socioeconomic status (SES) moderates the effect of equal-status contact. Staging intergroup contact online allows us to suppress—or reveal—real-world status cues that would be readily perceived in person[13]. Three weeks after contact, the positive effect of equal-status contact remained when informing participants about their paired partners' real-world SES.

Our primary contribution is to build on previous research on the contact hypothesis. We experimentally manipulate participants' status during contact and show that even brief collaborative interactions can have enduring positive effects. Previous experimental studies demonstrate that prolonged and intense contact between ethnic or religious groups, for instance, in sports leagues, can lessen discriminatory behaviours[14–17]. Shorter interventions have also reduced some kinds of non-political prejudice[18], but intergroup contact experiments that focus specifically on political prejudice have yielded mixed findings about effect persistence beyond treatment day[19–21]. In contrast to the present study, these and related approaches do not induce variation in participants' relative status while interacting[10].

Our study also contributes to research on affective polarization. Affective polarization is traditionally defined as the tendency of people to view opposing partisans negatively and copartisans positively[2,22]. Intolerant behaviour, our main dependent variable, is a key dimension of animosity towards outpartisans[2,23]. We demonstrate the importance of equal status as a condition for interpersonal contact to diminish partisan animosity. In addition, we provide proof of concept that online spaces for cross-partisan contact that put people on equal footing can diminish partisan animosity in an affectively polarized environment.

Our research demonstrates the reduction of intolerant behaviours by implementing actual contact, an approach that differs from interventions that do not use contact, but instead provide corrective information about outpartisans[24], give participants the opportunity to observe warm relations among opposing elites[25,26], or prime self-affirmation, empathy and other feelings[26–28]. Actual cross-partisan contact in our study also differs from interventions that simulate it by describing such contact in survey vignettes[29], simulating discussion environments[24] or having participants meditate and imagine contact[30].

We fielded our experiment in Mexico, where affective polarization has risen sharply in recent years[31,32]. The party system now features two poles, one in support of the incumbent party, Movimiento Regeneración Nacional (MORENA), and another that combines previous adversaries in opposition[33]. In 2021, the mean difference between inparty and outparty warmth on a 10-point feeling thermometer was 5.2, equalling the level of affective polarization in the United States in 2020. Partisan groups in Mexico divide over core political issues, including assessments of democracy and electoral integrity[34], and they harbour social animus towards one and other. Opposing partisans view each other as dishonest, intolerant and unpatriotic, they increasingly identify with different social classes[35,36], and they segregate into distinct social networks[37]. These divides are stoked regularly during President López Obrador's highly viewed daily morning press conferences that often include diatribes against the 'immoral' elites and the middle class, the media, and opposition parties and partisans[38]. Partisan sympathies in Mexico increasingly operate as a macrosocial identity, like they do in the United States, resulting in affective polarization marked by deep-seated intolerance that extends from political interactions to non-political social settings[1,5]. For instance, in a nationally representative survey, 79.8% of pro-MORENA respondents said that they would accept having a MORENA sympathizer live in their house; only 38.3% of anti-MORENA respondents felt similarly. (For details on all figures cited in this paragraph, see Supplementary Section 2.7 'Affective polarization in Mexico' and Supplementary Fig. 8.)

Results

Research design and sample

Our study includes a large number of participants. From an existing group of over 196,000 survey respondents in Mexico, we invited 3,120 individuals to join the study, in batches of several hundreds, at specific dates and times. Upon connecting, they were asked a question that we used to form pairs of citizens with opposing partisan sympathies (described below). We formed such pairs in the background while participants completed the remainder of a baseline survey. We then assigned pair-level treatments randomly, using blocking to improve statistical power (Methods). Pairs within a block were randomly assigned to equal status during contact (E, 780 pairs), unequal-status contact (U, 390 pairs with random assignment of pair members to Leader (U_L) or Follower (U_F)) or the no-contact control (C, 390 pairs). To probe robustness of equal-status contact to information about real-world status differences, we exposed a random subset of pairs in the equal-status condition E to information about their pair partner's SES (E_S, 390 pairs). The remaining pairs were not exposed to such information (E_N, 390 pairs).

Our main analysis sample consists of the 2,454 individuals (79% of those invited to the study) that remained after dropping those who did not complete the study and their paired partners. Attrition is statistically indistinguishable across treatment arms assigned to contact (E_S, E_N, U_F and U_L). Moreover, all experimental conditions including the no-contact control are well balanced on pre-treatment covariates, including turning out to vote in the 2018 presidential election, political interest, party identification, age, sex and SES, among others (Supplementary Section 1.1 'Balance').

Participants were asked to complete an endline survey directly after the intervention and a follow-up survey approximately 3 weeks later. Everyone received a participation fee upon completing the study. Additional incentives were provided within the study conditional on participant responses (Methods). All incentives were provided at the end of the study, no deception was used, and all protocols obtained Institutional Review Board (IRB) approval from The University of Texas at Austin and ITAM in Mexico City. We preregistered the trial at the Social Science Registry (https://www.socialscienceregistry.org/trials/8143). Tests of all preregistered hypotheses are reported in Supplementary Section 1.5 ('Additional analyses').

Pairing

We divided participants into pro- and anti-MORENA groups by asking them which party they would vote for if presidential elections were held today. We use this measure for the following reasons. First, research in other new democracies shows that vote choice performs better at measuring partisan sympathies than traditional party identification

questions that rely on prolonged exposure to the same party labels[39–41]. At the time of our study in 2021, MORENA had competed in just one presidential election in 2018. Second, although the two classification approaches (by vote choice versus party identification) create substantively similar groups, the large proportion of respondents who do not identify with any party means that the vote choice measure captures the pro- versus anti-MORENA cleavage for a much larger proportion of our sample. Supplementary Fig. 9 shows, using a separate 2018 survey, that views on core political issues such as evaluations of corruption, incumbent performance and the economy are virtually identical when classifying voters according to vote choice or partisan identification.

Pair-level intervention

After informing paired participants of their partner's partisanship (and, in the inequality condition, of whether they were designated Leader or Follower), we asked members of a pair to complete non-political tasks. Like other studies, we designed our tasks to enhance perceived commonality and foster collaboration. In the first task, participants were asked to decide whether Mexicans in general value friendship or professional success more highly. This task was designed to generate a connection quickly in the tradition of the 'fast friends' studies in social psychology[42] and prime participants' superordinate identity of being Mexican to enhance perceived commonality[43]. The second task further encouraged collaboration by asking participants to answer three trivia questions about Mexican popular culture. Light-hearted games like this one can increase positive attitudes towards outgroup members[44]. The tasks were accessible to participants independent of education, income, age, gender and political views. During the contact interaction, we held constant across contact treatment arms the presence of common goals and the incentive to collaborate, both of which Allport[6] hypothesized, and subsequent studies have shown, to enhance the effectiveness of intergroup contact[9,17]. A full description of the tasks appears in Supplementary Sections 2.3 ('Chat instructions') and 2.4 ('Control instructions').

Paired participants were invited to communicate in an open-ended manner with their partner while completing the tasks. Communication took place in an anonymous text-chat window on the same screen, powered by Chatter, a purpose-built application[19]. Figure 1 provides a flavour for the interface by showing screenshots of the first task and an example of several typical messages exchanged between participants assigned to contact under status equality. Each member of the pair entered responses to the tasks' questions individually and their responses could not be observed by their partner in any treatment condition. We informed participants that both members of the pair would qualify for entry into drawings, one for each task, if they provided answers to the values question and at least two of three correct answers to the trivia questions. The reward for the values task was the use of responses in teaching about Mexicans' values in universities in the United States and Mexico. The reward for the trivia task was a lump sum of cash points, roughly equivalent to 315 Mexican pesos or about US$15, that could be exchanged for goods in an online store.

We manipulated relative status by informing participants about their experimentally assigned status at the beginning of the first task and reminding them of it before the second task. In the equal-status treatment, participants were told that their respective answers would count equally. Specifically, one set of answers would be selected at random with equal probability to determine pair-level rewards. In the unequal-status treatment, participants were told that one member of the pair was randomly designated the 'Leader' and the other member the 'Follower', and only the Leader's answers would count for determining pair-level rewards. Individuals in the control condition completed the tasks and accrued rewards individually, without interpersonal contact. Collaboration between pair members lasted 10 min. Personal identifying information was never shown. In the unequal-status condition, chat handles read 'Leader' and 'Follower'.

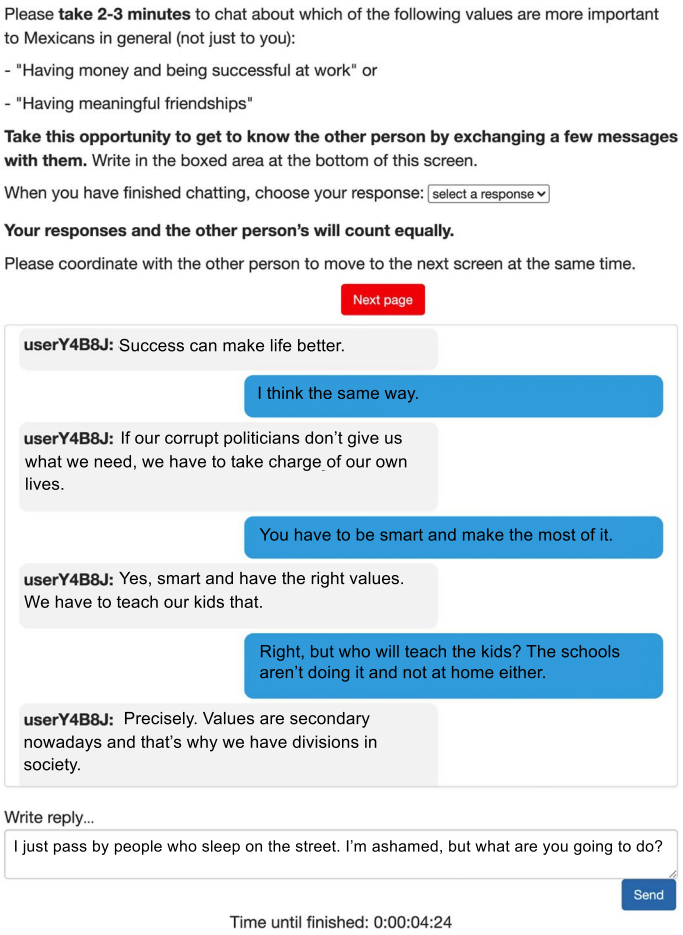

Fig. 1 | Chatter interface: sample chat. Note: This example pertains to the first task. The instructions were translated from Spanish by the authors. Chat contents show the spirit of a real conversation between a pair of participants assigned to the equal-status condition; however, because the researchers did not explicitly request permission to publish chat contents, we have altered the phrasing and expressed the conversation in English.

Our approach builds on recent work to manipulate equal status during intergroup contact, recognized as a difficult endeavour[45,46]. Lowe[17] experimentally manipulates payment equality in an Indian cricket league and finds that it does not undermine the beneficial effects of contact. Our design differs in that it directly manipulates relative status while maintaining payment equality across pair members, thereby holding constant incentives to collaborate. We discuss the complementary relationship between Lowe's research and ours in Supplementary Section 1.4 ('Main results').

Main outcome variables

We measured tolerance using incentivized behaviours. Sharing was measured through a dictator game where participants could choose to donate cash points, exchangeable for goods at an online store, to an anonymous participant with opposing partisan sympathies. Willingness to dialogue was measured as the response to an invitation to take part in a future online meeting with other participants, which we indicated would include opposing partisans and last 30 min. We aggregated the two measures into a standardized additive index of tolerant behaviour for expository ease, to improve statistical power by reducing measurement error, and to mitigate issues with multiple testing[47].

We selected these measures for four reasons. First, we sought to measure tolerance because it is central to Allport's predictions about the beneficial effects of contact[6]. Second, reducing intolerance is a key goal of interventions designed to ameliorate affective

https://doi.org/10.1038/s41562-024-02043-y

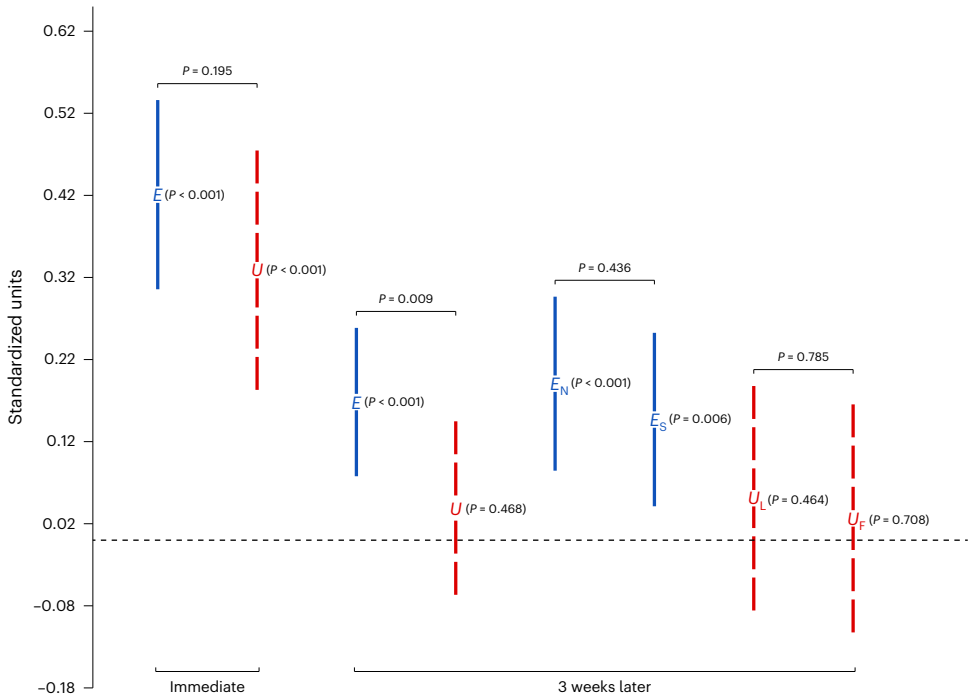

Fig. 2 | Index of tolerant behaviour 3 weeks after contact. Note: Point estimates of intent-to-treat effects are represented by treatment assignment indicators: blue E = equal status; red U = unequal status; blue E_N = equal status without revealing SES; blue E_S = equal status with SES revealed; red U_L = unequal status, assigned as Leader; red U_F = unequal status, assigned as Follower. Blue solid bars represent 95% CIs for estimates related to E conditions. Red dashed bars represent 95% CIs for estimates related to U conditions. The vertical axis is measured in standard deviations compared with the no-contact control condition. Inline P values compare each treatment arm to the no-contact control condition (two-sided). Horizontal P values correspond to difference-of-means tests between adjacent estimates (two-sided). Elements above the 'Immediate' label are computed with data from the endline survey, N = 1,550 (the difference with the full partnership sample size reported in Supplementary Table 1 is due to missing responses in the dependent variable). Elements above the '3 weeks later' label are computed with data from the follow-up survey, N = 2,454 (Supplementary Table 13).

polarization, and behavioural measures are increasingly used for this purpose[2–4,22,48]. Third, the specific tolerant behaviours we measure are vital to democracy. Democratic theorists view willingness to dialogue across partisan lines as key to problem solving and to elucidating a society's priorities[49,50]. In addition, sharing with outgroup members, for example, via taxation and redistribution, implies that people perceive the interests of others as legitimate[51]. Finally, incentivized behaviours are less susceptible to social desirability biases and experimenter demand effects than unincentivized survey questions because they make tolerant behaviour costly[52].

Estimates

Participants were attentive to the experience of interpersonal contact. Three weeks after treatment, 88% of participants assigned to a contact condition recalled having chatted; only 10% of those assigned to the no-contact control reported (erroneously) having chatted (equal status $\beta = 0.78$; $P < 0.001$; 95% CI, 0.74, 0.82; unequal status $\beta = 0.78$, $P < 0.001$; 95% CI, 0.73, 0.83; Supplementary Table 12).

Figure 2 shows intent-to-treat effects (Supplementary Tables 13 and 17). Immediately after treatment, assignment to intergroup contact under equal status increased the tolerant behaviour index by 0.42 standard deviations (σ) compared with no contact (leftmost estimate, $\beta = 0.42$; $P < 0.001$; 95% CI, 0.31, 0.54) and unequal-status contact increased it by 0.33σ (second estimate from the left, $\beta = 0.33$; $P < 0.001$; 95% CI, 0.18, 0.48; equal minus unequal status difference $P = 0.20$, 95% CI, −0.05, 0.23). Three weeks later, assignment to equal-status contact still increased the tolerant behaviour index by 0.17σ compared with no contact (third estimate from the left, $\beta = 0.17$; $P < 0.001$; 95% CI, 0.08, 0.26), while contact under unequal status could not be distinguished from the control (fourth estimate from the left, $\beta = 0.04$; $P = 0.47$; 95% CI, −0.07, 0.15). At the 3-week mark, the effect of assignment to equal-status contact was also 0.13σ larger than that of contact under unequal status (equal minus unequal status difference $P = 0.009$; 95% CI, 0.03, 0.23). The two rightmost estimates show that tolerant behaviour could not be statistically distinguished from the control for either Leaders or Followers in the unequal-status condition after 3 weeks. One implication of these findings is that, contrary to Allport's conjecture, equality may not be necessary to increase tolerance in the short run; however, consistent with his conjecture, equality may be necessary for the effect on tolerance to persist. We view the longer-term results as more relevant for policy[10] and potentially less prone to experimenter demand effects.

Participants in equal-status contact also experienced higher-quality interactions with opposing partisans. The number of phrases expressing agreement, such as 'you are right', 'yes' and 'I agree', were 10% more common under equal- versus unequal-status assignment (unequal status $\beta = -0.247$; $P = 0.036$; 95% CI, −0.48, −0.02), although this difference is not statistically significant when normalizing the number of agreement words by total words. The number of words in chat was more evenly distributed across members of pairs assigned to equal status compared with unequal status (unequal status $\beta = 170.34$; $P = 0.016$; 95% CI, 32.21, 308.47) (Supplementary Table 20).

The literature on the contact hypothesis in social psychology identifies three major categories of mechanisms through which high-quality contact might increase tolerance: learning that the outgroup is more similar to the ingroup than one thought, reduced anxiety about the outgroup, and perspective taking or empathy[7]. Although it is beyond the scope of our analysis to pin down a precise mechanism, suggestive findings can be found in Supplementary Tables 15 and 16 for chatting with an outparty stranger while waiting in line for a routine task (an indicator of anxiety), opinions of an outparty voter (an indicator of anxiety and perspective taking), and perceptions of shared values, outgroup honesty and outgroup intelligence.

We probed the robustness of our main findings by revealing participants' real-world SES to a random subset of pairs assigned to equal-status contact (treatment arm E_S). We elicited SES information before contact by asking participants to choose among five sets of images of house facades, kitchens and bedrooms corresponding to different socioeconomic strata, those that best represented their own homes (Supplementary Fig. 7). Exposure to real-world SES information increased participants' ability to correctly predict their paired partner's SES by 19% ($\beta = 0.07$; $P = 0.020$; 95% CI, 0.01, 0.13) (Supplementary Table 11).

Introducing information on real-world SES could in principle reinforce partisan animus among pair members with unequal SES, undercutting the potential for contact to enhance tolerant behaviour. The beneficial effect of equal-status contact, however, proved robust to the revelation of SES information. Figure 2 shows that the effects of E_N and E_S, compared with C, are statistically indistinguishable. The effect of equal-status assignment was also statistically indistinguishable across individuals whose real-world SES was higher, equal or lower than their partner's (Supplementary Tables 14 and 22). An alternative interpretation of this finding is that the procedure we used was not effective at memorably conveying information about a partner's SES.

Finally, we consider whether the lack of improvement in tolerant behaviour among participants in the unequal-status condition might have been driven by the displeasure of those randomly assigned to be Followers. However, there were no statistically significant differences between Leaders and Followers in their dictator-game donations to outparty participants or in their willingness to complete the follow-up survey 3 weeks after treatment (Supplementary Tables 2 and 13).

Our main focus is on the impact of contact on tolerant behaviours; nevertheless, we also report results for feeling thermometers and other attitudinal outcomes in Supplementary Section 2.9 ('Outcome variables'). Like other studies of intergroup contact[10,14,15,17], we do not find statistically significant effects on these attitudes at the 3-week mark (Supplementary Tables 15 and 17).

Discussion

Intensifying partisan polarization in many countries is straining democracy's moorings. Sympathizers of opposing parties frequently self-sort into different neighbourhoods, absorb news from different sources, and participate in different online social circles. When cross-partisan contact does occur, it is often brief and bitter, with intolerance exacerbated by social status differences. Nearly 200 years ago, Alexis de Tocqueville wrote that democracy thrives when citizens interact in the public square as equals[53]. Our design put a modern version of de Tocqueville's idea—and a decades-old conjecture that is central to contact theory—to the test by experimentally manipulating status within the interaction.

As it turns out, relative status in the interaction plays a key role in moderating the effects of intergroup contact on tolerant behaviours, consistent with Allport's view. When we place people in a situation of status equality, tolerant behaviours rise meaningfully and durably because of interpersonal contact. We find that revealing or withholding real-world SES makes no difference to the salutary effect of contact under equality.

We designed several features of our experiment to closely mimic online contact in everyday life. The chat app looks like other online interfaces, worked seamlessly across devices that participants routinely use such as computers, tablets and cell phones, and permitted participants to engage in chat wherever they preferred. In addition, we asked respondents to engage in collaborative tasks associated with natural topics of conversation. We thus believe that our intervention could be replicated and achieve similar results when staged online among other affectively polarized populations. Our study also has limitations. It is not clear whether our findings generalize to in-person contact, where status markers are readily perceived, or to contact among people who are not strangers and habitually interact under conditions of hierarchy. The durability of treatment effects beyond 3 weeks also remains an open question.

Our findings suggest practical ways of increasing mutual tolerance among opposing partisans. A version of intensive and costly in-person deliberation has recently been shown to improve cross-partisan understanding[54]. We believe that online spaces for cross-partisan contact that put people on equal footing can generate prosocial and democracy-supporting behaviour affordably and at large scale using a medium that is increasingly popular for political speech. Our results indicate that such spaces require only mild curation: conversations need not be orchestrated around political topics nor do opposing partisans have to be steered away from disagreement. Even under severe political polarization, enhancing tolerance of outgroups is within reach.

Methods

Ethical considerations

The study was approved before data collection by The University of Texas at Austin Institutional Review Board under exempt status (IRB ID STUDY00001126), with the determination that "this protocol meets the criteria for exemption from IRB review under 45 CFR 46.104 (2)(i) Tests, surveys, interviews, or observation (nonidentifiable)(3)(i)(A) Benign behavioural interventions (non-identifiable)". The study also adheres to the American Political Science Association's Principles and Guidance for Human Subjects Research.

To obtain informed and voluntary consent, all participants were shown an information screen before starting the baseline survey (pre-treatment). Participants were asked to indicate consent by clicking on an arrow on the screen to proceed with the study.

Participants were compensated by NetQuest in 'Korus', which they can use to buy gifts, enter raffles and/or make donations to NGOs. For instance, a ticket to the movies costs 95 Korus. A $1,000-Peso Amazon gift card costs 1,625 Korus. All participants who completed the brief recruitment survey were paid 4 Korus. Participants who completed their randomly assigned task and post-treatment survey, which took between 25 min and 45 min, were paid 100 Korus. Participants who returned for the 5 min follow-up survey were paid 10 Korus. Participants could also earn Korus based on their responses to some of the outcome measures, detailed in Supplementary Section 2.9 ('Outcome variables') and through raffles associated with some treatment conditions. We conducted those raffles within 1 week of the follow-up survey.

Sample recruitment

We recruited participants from within NetQuest's survey panel in Mexico. NetQuest, a private firm, has about 1.5 million panelists worldwide and over 196,000 across Mexico. Invitations were issued by email, in batches of several hundreds, to connect at a pre-specified time and date to conduct the actual study (further details on the logistics are provided below). A very small fraction of panelists who reported not being registered to vote were screened out. We asked NetQuest to make a special effort to recruit across the socioeconomic spectrum since lower-income deciles are under-represented in their pool of panelists. Our sample was slightly younger and more educated than the 2020 census population of adults, which is unsurprising for an online study (Supplementary Table 6). We maximized sample size given our budget. Our analysis sample size is larger than those used in other recent studies featuring dyadic chat-based experiments[20,21]. Data collection for the baseline survey, randomized component and the endline survey began on 24 August 2021 and continued until 1 September 2021. The follow-up survey was in the field approximately 3 weeks later depending on when the participant initially participated, from 17 September 2021 through 24 September 2021. Preregistration was completed on 26 August 2021.

Randomization procedure

Our experimental design uses blocked cluster randomization. Clusters consist of pairs of participants with opposing partisan sympathies.

Treatment is assigned at the pair level. Information on participants' partisan sympathies was gleaned from a survey question that asked which party the participant would vote for if the election for president were held today. If participants selected 'MORENA', we deemed them pro-incumbent. If participants selected any other political party, we deemed them anti-incumbent. All pairs consisted of one pro- and one anti-incumbent individual thus defined.

To increase statistical power, we used information about participants gathered during the baseline survey to create pairs and organize them into blocks. The variables we used for blocking include feeling thermometers towards the three largest political parties (PRI, PAN and MORENA) and towards the incumbent president, as well as SES. We validate these measures and provide more details on the composition of pro- and anti-incumbent partisan sympathizers in Supplementary Section 2.5 (Supplementary Tables 7, 8 and 10).

The algorithm we used to create blocked partnerships is illustrated in Supplementary Fig. 1. First, we divided participants into pro- and anti-incumbent partisan sympathizers based on the vote choice question referenced above. Second, within these partisan groups, using an optimal Greedy algorithm, we selected four participants to form clusters of similar people by minimizing Mahalanobis distance (computed using feeling thermometers towards people who identify with the three largest parties, feelings towards AMLO and the picture-based measure of SES). We used the blockTools software to implement this step[55]. Third, we randomly assigned the sets of four similar partisan sympathizers to an analogous set of four cross-partisans. This creates a block. Fourth, within each block, we randomly assigned each participant to a cross-partisan partner, thus forming four pairs. We also did this for the control group, even though the control condition entailed no interaction between members of a pair. Within each block, we then randomly assigned pairs to experimental conditions (C, U, E_N, E_S) using the randomizr R package[56]. For pairs assigned to the U condition, we additionally randomly assigned participants to be either the Leader (U_L) or the Follower (U_F).

Randomization achieved balance. Supplementary Table 3 shows balance across treatment conditions on blocking covariates as well as other important covariates measured before treatment. We also assess balance in the main sample, which includes only participants that completed the follow-up survey (fielded approximately 3 weeks after treatment). Supplementary Table 5 shows that there is no differential attrition between equal (E) versus unequal status (U) treatment arms. Attrition is slightly greater for participants assigned to the contact conditions compared with the no-contact control C (about 3% for E and 1% for U), but all treatment arms and the no-contact control are balanced on observables (Supplementary Section 1.2 'Attrition'). Finally, the sample that completed the survey fielded immediately after treatment is also balanced on pre-treatment covariates (Supplementary Table 4).

Treatment assignment and data collection were automated. The research team was not blind to treatment assignment during data collection and analysis.

Online conversation logistics

This section describes the procedure we used to coordinate pairwise online contact sessions.

The chat portion of the experiment took place on 24–26 August 2021 and 30 August to 1 September 2021. On each of these 6 days, at a pre-specified time of day, participants followed a link that directed them to the study. They completed a baseline survey, were randomized to partners and to treatments as described above, completed the synchronous chat (except for those in pairs assigned to the no-contact control condition, who did not chat) and associated tasks, and took the endline survey, all within 40 min.

The day before each study day, we sent out an invitation email that included a link to a short screening survey to several hundred panelists. That survey alerted them to the fact that they might receive an invitation the next morning to join the study in the afternoon. We issued invitations only to those who stated that they were eligible to vote and available at the time of the study. We aimed to invite about half MORENA supporters and half supporters of other political parties, since anyone who logged on but could not be paired could not participate in the study (recall that treatment assignment was conducted at the pair level). These dropped participants would still have spent time answering the baseline questionnaire and would have been paid a participation fee, wasting project resources.

The invitation email sent the day before the study that contained the link for the screening survey also included a link for the main study. Panelists were asked to log on at a specific time in the afternoon, to participate in the study. They were informed that they would not be able to participate if they logged on more than 5 min after the appointed time. Upon clicking on the link, participants were directed to the baseline survey, which was programmed on Qualtrics. At the start of the survey, participants were asked the questions that would later be used by our algorithm as blocking covariates, other than the vote choice question that we had asked the day before in the screening survey. Then, while participants completed the remainder of the baseline survey, we downloaded the blocking covariate data from Qualtrics to feed it to the algorithm that would implement our experimental design (that is, blocking, pairing participants and randomizing treatments). Simultaneously, we used this window of time to create online chatrooms for each pair. To do this, we utilized the Chatter software, which has an API that researchers can use to create users, create chatrooms with specific instructions (that is, our treatments) and assign users to chatrooms.

Fewer than 3% of participants that logged on were dropped from the study because they did not complete the pairing questions that appeared towards the start of the baseline survey within 13 min of the appointed start time (that is, 8 min after the end of the 5 min grace period) or, in a very few cases, because they failed the simple attention-check question in the baseline survey.

At the 15 min mark after the appointed time, participants could advance to their pre-assigned chatroom with their pre-assigned partner (participants who completed the baseline before that moment or before their partner arrived were asked to wait so that both members of a pair would begin the chat section of the study simultaneously).

Participants then chatted and completed the tasks outlined in the 'Treatment conditions' section (Supplementary Section 2.2) while a visible 8 min timer counted down from the moment that both members of the pair entered the chatroom. When the timer ran out, participants could advance to the endline survey, also programmed on Qualtrics.

Approximately 3 weeks later, all participants who were assigned to a pair at baseline were invited to complete a follow-up survey. The follow-up survey was in the field during 17–24 September 2021, with 75% of participants completing the survey the day it was distributed. Therefore, a majority of participants completed the follow-up 2.5 to 3.5 weeks after treatment.

Estimation

Our main results are estimates of intent-to-treat effects based on the following equation:

$$Y_{i,p(i),k(p)} = \alpha + \beta_E E_{i,p(i)} + \beta_U U_{i,p(i)} + \boldsymbol{\gamma} X_{p(i)} + \theta_{k(i)} + \varepsilon_{i,p(i),k(p)}, \quad (1)$$

where i indexes individuals, p indexes the pair to which the individual belongs, and k indexes the block to which the pair p belongs (reflecting the block randomization). $Y_{i,p(i),k(p)}$ is the outcome variable of interest for individual i in block k (for instance, the tolerant behaviour index). $E_{p(i)}$ is an indicator for assignment of pair p to the equal-status condition, and $U_{p(i)}$ indicates assignment to unequal status. The control group $C_{p(i)}$ is the omitted category; therefore, average outcomes for the control group are captured by α.

Block dummies $\theta_{k(i)}$ reflect the blocked randomization design. $X_{p(i)}$ is a matrix of pre-treatment covariates, included to increase statistical power. These covariates include the outcome variable $Y_{i,p(i),k(p)}$ measured at baseline and an indicator variable (MORENA_i) that takes the value of 1 if individual i is a MORENA supporter. $\boldsymbol{\gamma}$ is a vector of regression coefficients; ε denotes residuals. Standard errors are robust and clustered at the pair level.

Coefficient β_E estimates the intent-to-treat effect of pair-level assignment to equal-status contact (E) compared with no contact (C). Similarly, β_U estimates the intent-to-treat effect of assignment to unequal-status contact (U) compared with no contact (C). While the effects of contact versus no contact are of interest and constitute the focus of most of the relevant empirical literature, our primary focus is testing whether, conditional on contact taking place, contact under equal status is more effective than contact under unequal status. This requires testing β_E against β_U. We perform these hypotheses tests and report the P values in the regression tables.

We run additional specifications to estimate the effects of the randomizations we performed within the E and U conditions. These specifications are represented by equation (2):

$$\begin{aligned} Y_{i,p(i),k(p)} = \alpha + \beta_{E_N} E_{N_{i,p(i)}} + \beta_{E_S} E_{S_{i,p(i)}} + \beta_{U_L} U_{L_{i,p(i)}} \\ + \beta_{U_F} U_{F_{i,p(i)}} + \boldsymbol{\gamma} X_{p(i)} + \theta_{k(i)} + \varepsilon_{i,p(i),k(p).} \end{aligned} \quad (2)$$

As explained in the main text, a random subset of pairs assigned to E were assigned to information revelation about their partner's SES (E_S). In that condition, each individual in the pair learned which of the five sets of pictures (of house rooms and facades) was selected by their partner. The rest of the individuals in pairs assigned to E (denoted E_N) were not provided with information about their partner's SES. Moreover, individuals within a pair assigned to U were respectively randomized into the roles Leader (U_L) or Follower (U_F). The variables $E_{N_{i,p(i)}}$, $E_{S_{i,p(i)}}$, $U_{L_{i,p(i)}}$ and $U_{F_{i,p(i)}}$, respectively, indicate assignment of a pair to E_N and E_S, and assignment of an individual to U_L and U_F.

The effect of revealing real-world SES within a pair is estimated by comparing the coefficients β_{E_N} versus β_{E_S}. The effect of being a Leader versus a Follower (within a pair assigned to U) is obtained by comparing the coefficients β_{U_L} versus β_{U_F}.

We estimate all models using STATA 17 and the following packages: unique, egenmore, blindschemes, binscatter, leebounds, reghdfe, outreg2 and ftools. To measure features from the chats, we use the following packages in R: syuzhet, tm, SnowballC, wordcloud, RColorBrewer, ggplot2, rlang, dplyr and tidyr.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The original data collected for this study that are necessary to replicate all results in the paper and Supplementary Information are available at https://osf.io/f7bzy/. Owing to human subjects protections, the raw text from the chats is not available, but all text-based measures needed to replicate analyses are included in the public dataset. The paper and Supplementary Information also make use of publicly available datasets[57–63].

Code availability

The Stata and R code necessary to replicate all results[64–67] in the paper and Supplementary Information is available at https://osf.io/f7bzy/.

References

1. Carlin, R. E. & Love, G. J. Political competition, partisanship and interpersonal trust in electoral democracies. *Br. J. Political Sci.* **48**, 115–139 (2018).
2. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. J. The origins and consequences of affective polarization in the United States. *Annu. Rev. Political Sci.* **22**, 129–146 (2019).
3. Finkel, E. J. et al. Political sectarianism in America. *Science* **370**, 533–536 (2020).
4. Baldassarri, D. & Page, S. E. The emergence and perils of polarization. *Proc. Natl Acad. Sci. USA* **118**, 1–7 (2021).
5. Mason, L. *Uncivil Agreement: How Politics Became Our Identity* (Univ. Chicago Press, 2018).
6. Allport, G. *The Nature of Prejudice* (Addison-Wesley, 1954).
7. Pettigrew, T. F. & Tropp, L. R. How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *Eur. J. Soc. Psychol.* **38**, 922–934 (2008).
8. Enos, R. D. Causal effect of intergroup contact on exclusionary attitudes. *Proc. Natl Acad. Sci. USA* **111**, 3699–3704 (2014).
9. Pettigrew, T. F. & Tropp, L. R. A meta-analytic test of intergroup contact theory. *J. Pers. Soc. Psychol.* **90**, 751 (2006).
10. Paluck, E. L., Green, S. A. & Green, D. P. The contact hypothesis re-evaluated. *Behav. Public Policy* **3**, 129–158 (2019).
11. Paluck, E. L., Porat, R., Clark, C. S. & Green, D. P. Prejudice reduction: progress and challenges. *Annu. Rev. Psychol.* **72**, 533–560 (2021).
12. Fast, N. J., Halevy, N. & Galinsky, A. D. The destructive nature of power without status. *J. Exp. Soc. Psychol.* **48**, 391–394 (2012).
13. Desmichel, P. & Rucker, D. D. Social-rank cues: decoding rank from physical characteristics, behaviors, and possessions. *Curr. Opin. Psychol.* **43**, 79–84 (2022).
14. Scacco, A. & Warren, S. S. Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria. *Am. Political Sci. Rev.* **112**, 654–677 (2018).
15. Mousa, S. Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science* **369**, 866–870 (2020).
16. Bursztyn, L., Chaney, T., Hassan, T. A. & Rao, A. The immigrant next door. *Am. Econ. Rev.* **114**, 348–384 (2024).
17. Lowe, M. Types of contact: a field experiment on collaborative and adversarial caste integration. *Am. Econ. Rev.* **111**, 1807–1844 (2021).
18. Broockman, D. & Kalla, J. Durably reducing transphobia: a field experiment on door-to-door canvassing. *Science* **352**, 220–224 (2016).
19. Rossiter, E. *The Similar and Distinct Effects of Political and Non-Political Conversation on Affective Polarization* Working Paper (Univ. Notre Dame, 2023).
20. Rossiter, E. L. & Carlson, T. N. Cross-partisan conversation reduced affective polarization for republicans and democrats even after the contentious 2020 election. *J. Politics* **86**, 1608–1612 (2024).
21. Santoro, E. & Broockman, D. E. The promise and pitfalls of cross-partisan conversations for reducing affective polarization: evidence from randomized experiments. *Sci. Adv.* **8**, eabn5515 (2022).
22. Iyengar, S. & Westwood, S. J. Fear and loathing across party lines: new evidence on group polarization. *Am. J. Political Sci.* **59**, 690–707 (2015).
23. Hartman, R. et al. Interventions to reduce partisan animosity. *Nat. Hum. Behav.* **6**, 1194–1205 (2022).
24. Voelkel, J. G., Ren, D. & Brandt, M. J. Inclusion reduces political prejudice. *J. Exp. Soc. Psychol.* **95**, 104149 (2021).
25. Huddy, L. & Yair, O. Reducing affective polarization: warm group relations or policy compromise? *Political Psychol.* **42**, 291–309 (2021).
26. Voelkel, J. G. et al. Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nat. Hum. Behav.* **7**, 55–64 (2023).

27. Levendusky, M. S. When efforts to depolarize the electorate fail. *Public Opin. Q.* **82**, 583–592 (2018).
28. Santos, L. A., Voelkel, J. G., Willer, R. & Zaki, J. Belief in the utility of cross-partisan empathy reduces partisan animosity and facilitates political persuasion. *Psychol. Sci.* **33**, 1557–1573 (2022).
29. Wojcieszak, M. & Warner, B. R. Can interparty contact reduce affective polarization? A systematic test of different forms of intergroup contact. *Political Commun.* **37**, 789–811 (2020).
30. Simonsson, O., Narayanan, J. & Marks, J. Love thy (partisan) neighbor: brief befriending meditation reduces affective polarization. *Group Process. Intergr. Relat.* **25**, 1577–1593 (2022).
31. Moreno, A. Polarización presidencial. *El Financiero* https://www.elfinanciero.com.mx/opinion/alejandro-moreno/polarizacion-presidencial/ (4 September 2020).
32. Cornejo, R. C. Same scandal, different interpretations: politics of corruption, anger, and partisan bias in Mexico. *J. Elect. Public Opin. Parties* **33**, 497–518 (2023).
33. Cornejo, R. C. The AMLO voter: affective polarization and the rise of the Left in Mexico. *J. Politics Lat. Am.* **15**, 96–112 (2023).
34. Hernández-Huerta, V. & Cantú, F. Public distrust in disputed elections: evidence from Latin America. *Br. J. Political Sci.* **52**, 1923–1930 (2022).
35. Sánchez-Talanquer, M. & Greene, K. F. Is Mexico falling into the authoritarian trap? *J. Democr.* **32**, 56–71 (2021).
36. Moreno, A. Lucha de clases, en modo electoral. *El Financiero* https://www.elfinanciero.com.mx/opinion/alejandro-moreno/2022/06/24/lucha-de-clases-en-modo-electoral/ (24 June 2022).
37. Signa Lab ITESO. México 2019: la disputa por la interpretación. https://signalab.iteso.mx/informes/informe_disputa-polarizacion_01.html (23 May 2019).
38. Sarsfield, R. Between the pure people and the corrupt elite. Populist storytelling and affective polarization in social media in Mexico. *Rev. Mex. Opin. Pública* **35**, 13–34 (2023).
39. Brader, T. & Tucker, J. A. The emergence of mass partisanship in Russia, 1993-1996. *Am. J. Political Sci.* **45**, 69–83 (2001).
40. Dinas, E. Does choice bring loyalty? Electoral participation and the development of party identification. *Am. J. Political Sci.* **58**, 449–465 (2014).
41. Baker, A., Ames, B. & Rennó, L. *Persuasive Peers: Social Communication and Voting in Latin America* (Princeton Univ. Press, 2020).
42. Aron, A., Melinat, E., Aron, E. N., Vallone, R. D. & Bator, R. J. The experimental generation of interpersonal closeness: a procedure and some preliminary findings. *Pers. Soc. Psychol. Bull.* **23**, 363–377 (1997).
43. Levendusky, M. S. Americans, not partisans: can priming American national identity reduce affective polarization? *J. Politics* **80**, 59–70 (2018).
44. Rajadesingan, A. et al. GuesSync!: an online casual game to reduce affective polarization. *Proc. ACM Hum.-Comput. Interact.* **7**, 1–33 (2023).
45. Riordan, C. Equal-status interracial contact: a review and revision of the concept. *Int. J. Intercult. Relat.* **2**, 161–185 (1978).
46. Busby, E. *Should You Stay Away from Strangers? Experiments on the Political Consequences of Intergroup Contact* (Cambridge Univ. Press, 2021).
47. Anderson, M. L. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *J. Am. Stat. Assoc.* **103**, 1481–1495 (2008).
48. Carlin, R. E. & Love, G. J. The politics of interpersonal trust and reciprocity: an experimental approach. *Political Behav.* **35**, 43–63 (2013).
49. Habermas, J. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (MIT Press, 1991).
50. Barber, B. *Strong Democracy: Participatory Politics for a New Age* (Univ. California Press, 2003).
51. Alesina, A. & Giuliano, P. in *Handbook of Social Economics* 1st ed. (eds Benhabib, J., Bisin, A. & Jackson, M. O.) Vol. 1, 93–131 (Elsevier, 2011).
52. Bauer, M., Chytilová, J. & Miguel, E. Using survey questions to measure preferences: lessons from an experimental validation in Kenya. *Eur. Econ. Rev.* **127**, 103493 (2020).
53. de Tocqueville, A. *Democracy in America* Vols. I and II (Read Books Ltd., 2015).
54. Fishkin, J., Siu, A., Diamond, L. & Bradburn, N. Is deliberation an antidote to extreme partisan polarization? Reflections on "America in One Room". *Am. Political Sci. Rev.* **115**, 1464–1481 (2021).
55. Moore, R. T. Multivariate continuous blocking to improve political science experiments. *Political Anal.* **20**, 460–479 (2012).
56. Coppock, A. randomizr: Easy-to-Use Tools for Common Forms of Random Assignment and Sampling. R package version 1.0.0. https://declaredesign.org/r/randomizr/ (2024).
57. Comparative Study of Electoral Systems (CSES). CSES Module 5 Full Release [dataset and documentation]. 25 July 2023 version. https://doi.org/10.7804/cses.module5.2023-07-25 (2023).
58. American National Election Studies (ANES). ANES 2020 Time Series Study Full Release [dataset and documentation]. 19 July 2021 version. www.electionstudies.org (2021).
59. Instituto Nacional Electoral (INE). Sistema de Consulta de la Estadística de las Elecciones (2018 Presidential Election Results by Municipality). https://siceen21.ine.mx/busqueda/Presidencia%20de%20la%20Republica/2/1/2018 (2018).
60. Instituto Nacional Electoral (INE). Sistema de Consulta de la Estadística de las Elecciones (2021 Presidential Election Results by Municipality). https://siceen21.ine.mx/busqueda/Diputaciones%20de%20MR/5/1/2021/2 (2021).
61. Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL). Pobreza a Nivel Municipio 2010–2020. https://www.coneval.org.mx/Medicion/Paginas/Pobreza-municipio-2010-2020.aspx (2021).
62. Instituto Nacional de Estadística y Geografía (INEGI). Censo de Población y Vivienda 2020. https://www.inegi.org.mx/programas/ccpv/2020/#datos_abiertos (2021).
63. Greene, K., Simpser, A., Ponce, A., Parás, P. & López, C. *Mexico 2018 Elections and Quality of Democracy (EQD) Survey* (Univ. Texas at Austin [distributor], 2018).
64. Mohammad, S. & Turney, P. Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. In *Proc. NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* (eds Inkpen, D. & Strapparava, S.) 26–34 (Association for Computational Linguistics, 2010); https://aclanthology.org/W10-0204
65. Mohammad, S. M. & Turney, P. D. Crowdsourcing a word-emotion association lexicon. *Comput. Intell.* **29**, 436–465 (2013).
66. Lee, D. S. Training, wages, and sample selection: estimating sharp bounds on treatment effects. *Rev. Econ. Stud.* **76**, 1071–1102 (2009).
67. Gapminder. Dollar Street, Families. https://www.gapminder.org/dollar-street

Acknowledgements

We thank M. P. de León, R. G. Tellez, R. Santa Maria and J. S. Córdova for excellent research assistance, as well as helpful comments from C. Adida, J. S. Arce, S. Badrinathan, R. Blair, D. Broockman, S. Chauchard, J. R. Enríquez, J. Gerring, J. Gottlieb, E. Kramon, M. Lowe, L. Mason, G. Munck, D. Nielson, M. Schenk, A. Scacco, C. Woods and participants in the General Seminar at the Institute for Advanced Study of Toulouse and a panel at the 2023 American Political Science Association Annual Meeting, and the anonymous referees. K.F.G. and A.S. gratefully

acknowledge funding from ConTex (2018-63), a consortium of Mexico's National Council of Science and Technology (CONACYT) and the University of Texas System. A.S. gratefully acknowledges funding from the French Agence Nationale de la Recherche under the Investissement d'Avenir programme (ANR-17EURE-0010). E.L.R. gratefully acknowledges funding from the Francis and Kathleen Rooney Center for the Study of American Democracy, University of Notre Dame. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper.

Author contributions

All authors contributed equally. Author names are listed in alphabetical order.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-024-02043-y.

Correspondence and requests for materials should be addressed to Alberto Simpser.

Peer review information *Nature Human Behaviour* thanks Matt Lowe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Alberto Simpser

Last updated by author(s): Jul 9, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Survey data were collected via Qualtrics. We also utilized a custom-made, original software called Chatter to implement the contact/ conversation portion of the study. The only data collected from Chatter were the transcripts of the chats.
Data analysis	We used STATA Version 17 to analyze the data and produce tables and figures. We used the following packages to do so: unique (10.1), egenmore (3.4.2), blindschemes (no versioning), binscatter (7.02), leebounds (1.5 2013-07-17), reghdfe (6.12.3), outreg2 (2.3.2), ftools (2.49.1). The data and code to replicate all tables and figures in the article and Supplementary Information have been deposited in a public repository at https://osf.io/f7bzy/. We used R to randomize participants to experimental conditions. We used the following R packages to do so: blockTools (0.6-3) and randomizr (0.8.0). This code is also available in the replication archive at https://osf.io/f7bzy/. Finally, we used R to measure features of the chats from the text. We used the following R packages to do so: syuzhet (1.0.6), tm (0.7.13), SnowballC (0.7.1), wordcloud (2.6), RColorBrewer (1.1.3), ggplot2 (3.5.1), rlang (1.1.4), dplyr (1.1.4), and tidyr (1.3.1). This code, and the measures it generates, are also available in the replication archive at https://osf.io/f7bzy/.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The original data collected for this study that is necessary to replicate all results in the paper and Supplementary Information are available at https://osf.io/f7bzy/. Due to human subjects protections, the raw text from the chats are not available, but all text-based measures needed to replicate analyses are included in the public dataset. The article and Supplementary Information also make use of the following publicly available datasets:

The Comparative Study of Electoral Systems (CSES). CSES Module 5 Full Release [dataset and documentation]. URL doi:10.7804/cses.module5.2023-07-25. July 25, 2023 version.

American National Election Studies (ANES). ANES 2020 Time Series Study Full Release [dataset and documentation] (2021). URL www.electionstudies.org. July 19, 2021 version.

National Electoral Institute (INE). 2018 Presidential Election Results by Municipality. URL https://siceen21.ine.mx/busqueda/Presidencia%20de%20la%20Republica/2/1/2018. Accessed August 12, 2021.

National Electoral Institute (INE). 2021 Presidential Election Results by Municipality. URL https://siceen21.ine.mx/busqueda/Diputaciones%20de%20MR/5/1/2021/2. Accessed August 12, 2021.

National Council for the Evaluation of Social Development Policy (CONEVAL). Poverty at the Municipal Level 2010-2020. URL https://www.coneval.org.mx/Medicion/Paginas/Pobreza-municipio-2010-2020.aspx. Accessed August 12, 2021.

National Institute of Statistical and Geographic Information (INEGI). 2020 Census of Population and Housing. URL https://www.inegi.org.mx/programas/ccpv/2020/#datos_abiertos. Accessed August 12, 2021.

Greene, Kenneth, Alberto Simpser, Alejandro Ponce, Pablo Parás, and Carlos López. Mexico 2018 Elections and Quality of Democracy (EQD) Survey (2018). Austin, TX: University of Texas at Austin [distributor].

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex was determined via self-report. The variable was used to check for pretreatment covariate balance across experimental treatment branches (supplementary materials Tables T-3, T-4, T-5).
Reporting on race, ethnicity, or other socially relevant groupings	We collected information on socioeconomic status (SES) to test whether SES moderated the causal effect of equal-status contact on tolerance. We used two SES variables. The first variable was provided by Netquest, the panel provider, for every participant. The categories of that variable are standard categories used in marketing studies in Mexico. The second variable was collected by the research team through a survey question that asked respondents to choose, from within 5 sets of pictures of homes, which one most resembled theirs (supplementary materials Figure F-7). The latter SES variable is the one we used in the main analysis (Figure 2 in the paper). The former SES variable was used primarily to check for pretreatment covariate balance across experimental treatment branches (supplementary materials Tables T-3, T-4, T-5), as well as to validate the latter SES variable (supplementary materials Table T-7).
Population characteristics	Population characteristics of our sample are described in supplementary materials Table T-6 and compared to corresponding 2020 Census data. Fraction of males was 45% in our sample and 48% in the Census. Fraction aged 30-39 was 30% in our sample and 21% in the Census. Fraction who completed high school was 19% in our sample and 20% in the Census. Fraction who completed university was 39% in our sample and 18% in the Census. Additional descriptives are provided in the balance tables (supplementary materials tables T-3, T-4, T-5).
Recruitment	Participants were recruited via Netquest, a survey panel provider. The research team at no point received names, addresses, email addresses, or other personal identifying information.
Ethics oversight	Project was approved by the IRBs at University of Texas at Austin and at ITAM.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study was a randomized experiment. The data are quantitative.
Research sample	The research sample was Mexican citizens from the Netquest survey panel. Characteristics of our sample are described in supplementary materials Table T-6 and compared to corresponding 2020 Census data. Fraction of males was 45% in our sample and 48% in the Census. Fraction aged 30-39 was 30% in our sample and 21% in the Census. Fraction who completed high school was 19% in our sample and 20% in the Census. Fraction who completed university was 39% in our sample and 18% in the Census. Additional descriptives are provided in the balance tables (supplementary materials tables T-3, T-4, T-5). The sample was a convenience sample. A convenience sample was needed to have a sufficient sample size because this study involved difficult-to-coordinate synchronous conversations.
Sampling strategy	The sampling strategy aimed to recruit an equal ratio of MORENA and non-MORENA supporters from the Netquest panel, as our design necessitated equal numbers of participants from these populations (one MORENA and one non-MORENA supporter were paired together in the design). In our preregistration document posted to the AEA RCT Registry, we stated samples sizes we would randomize to each treatment arm -- 350 pairs each. The control condition would have 700 individuals. We believed this would afford sufficient power.
Data collection	The data collection procedure took place entirely online. To obtain informed and voluntary consent, all participants were shown an information screen prior to starting the baseline survey (pre-treatment). Participants were asked to indicate consent by clicking on an arrow on the screen to proceed with the study. The consent screen as well as the baseline survey, the chat-based treatment, the endline survey, and the followup survey were all completed by participants alone via their electronic devices with no contact with the researchers or other participants. The researchers were not blinded to the study hypotheses nor the experimental conditions when collecting or analyzing the data.
Timing	Data collection for the baseline survey, randomized component, and the endline survey began on August 24, 2021 and continued until September 1, 2021. The followup survey was in the field approximately three weeks later depending on when the participant initially participated, from September 17, 2021 through September 24, 2021.
Data exclusions	N=3120 people were randomized in the design and N=3015 completed the pre-treatment baseline survey. (See next section for discussion of non-participation post-randomization). In the analyses, we included only full partnerships where both participants participated in the chat. For the main results in Figure 2 in paper, we used complete pairs at the endline (complete defined by both partners completed the measures in the Intolerant Behavior index (N=1550)) responses to estimate short-run treatment effects. We did not exclude participants for other reasons. For the long-run treatment effects (three-weeks later), we again used the same definition of complete pairs (N=2454). This decision rule was used throughout all analyses in the SI, as well. Finally, participants with missing data for any of the variables used in any given analysis were automatically dropped from that specific analysis (i.e., we did not impute missing data).
Non-participation	For the main results based on the followup survey (three weeks after treatment) we lose 19% of the baseline sample (i.e., the N=3015 participants who completed the baseline survey). For the analysis of effects immediately following treatment (endline survey), we lose 42% of those who completed the baseline survey. These exclusions are due to individual attrition and to the fact that we used only complete pairs.
Randomization	We used blocked cluster randomization to assign pairs (clusters) of participants to experimental arms. Each block consisted of four similar pairs (clusters) on dimensions important to the outcome (ALMO feeling thermometer, PAN feeling thermometer, PRI feeling thermometer, MORENA feeling thermometer, and picture-based SES measure). After creating blocks, we randomly assigned the four experimental conditions at the pair-level within each block.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.