

# Contact in the Classroom: Can Virtual Exchanges Build Tolerance?

Salma Mousa\*

May 30, 2020

## Abstract

Can virtual contact reduce prejudice? Despite extensive research on intergroup contact, and the practical benefits of virtual exchanges, we have little causal evidence on the effects of virtual contact. I shed light on this question using a natural experiment among  $n = 508$  primarily North American students aged 6 – 11, roughly half of whom are assigned to classrooms that conduct video exchanges with peers in another part of the world. Compared with the control group, whose attitudes worsened over time, assignment to these classrooms made students 11 – 25% more likely to perceive commonalities with foreigners and ethnic outgroups, 12% more likely to describe these groups positively, 14% more likely to practice empathy, and 10% more likely to report that their friends would accept an intergroup friendship, suggesting that contact boosts tolerant social norms and perspective-taking. These results are driven, however, by exchanges that bring together groups with some ethnic overlap, suggesting that social distance moderates the effects of contact. The findings show that virtual contact, bundled with a teacher committed to diversity, can stabilize intergroup attitudes in the face of rising prejudice outside the classroom.

\*Department of Political Science, Stanford University; 616 Serra St., Encina Hall West, Palo Alto, CA 94305, USA. smousa@stanford.edu. I thank Chelsea Donaldson, Verónica Vázquez Ugalde, and Todd Hall for their excellent research management. I thank Lisa Blaydes, Jens Hainmueller, David Laitin, and Jeremy Weinstein for helpful feedback. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of Empatico.

American neighborhoods and schools are becoming increasingly segregated (Orfield et al. 2014; Bayer, McMillan and Reuben 2004). This residential sorting has lowered the frequency and quality of contact across social lines (Eric Oliver and Wong 2003), and correlates with a range of negative outcomes: from prejudice and discrimination toward ethnic minorities (Tropp and Prenovost 2008), and support for far-right parties and policies (Clayton, Ferwerda and Horiuchi 2018; Green et al. 2016), to weakened linkages needed for collective action (Habyarimana et al. 2007) and social capital (Putnam 2007), which themselves are key determinants of public goods provision (Alesina, Baqir and Easterly 1999; Putnam 1993) and economic growth (Rothstein 2005). Positive intergroup contact has helped to mitigate these effects in myriad contexts (Barnhardt 2009; Rao et al. 2013; Lowe 2020; Carrell, Hoekstra and West 2015), although more more causal evidence is needed (Paluck, Green and Green 2017), especially from classroom settings (Scacco and Warren 2018). Even if contact does reduce prejudice, however, there is a question of whether this approach is sustainable. If people are exposed to the same social and structural barriers that led to segregation in the first place, then “our current policies for encouraging contact may not be enough, even for those people the policies actually reach” (Enos 2014, p. 249).

One tool seems particularly well-suited to overcoming the structural barriers to intergroup contact: virtual interactions. Regardless of how segregated one’s neighborhood is, or the accessibility of public transportation, or the other social and economic costs associated with traditional contact, an internet connection is enough to facilitate an interaction across group lines. Should we expect virtual contact to improve intergroup relations? Meta-analytic tests conclude that contact “typically reduces prejudice” but highlight the lack of evidence on how different types of contact (e.g. in-person vs. virtual) can shape these effects (Pettigrew and Tropp 2006; Paluck, Green and Green 2017). On one hand, non-verbal cues, mirroring behaviors, and physical touch are known to form the micro-foundations of empathy in humans, but are either weak or absent from virtual interactions (Mehrabian 2017; Rizzolatti and Craighero 2004; Schroeder et al. 2014). On the other hand, indirect contact, typically in the form of fictional narratives, online role-playing games, and perspective-taking exercises, can build empathy in adults and children (Cameron and Rutland 2006; Alvídrez et al. 2015; Stathi et al. 2014; Todd and Galinsky 2014; Simonivitz, Kezdi and Kardos 2017; Adida, Lo and Platas 2018). These interventions involve fictional outgroup members and lack interaction, however. The effectiveness of virtual contact thus remains an open question.

This study is among the first causal tests of virtual contact. I leverage variation in exposure to virtual contact among  $n = 508$  students in 42 primarily North American classrooms, 75% of which are located in majority white, suburban, right-leaning enclaves. The ability to credibly identify the causal impact of such exposure comes from

two sources. First, parents are unaware of the virtual exchange program before the academic year, students do not choose their homeroom teacher, and over 95% of students in the sample are of “traditional learning ability,” precluding them from being hand-picked for certain classrooms based on special learning needs. Second, I mitigate the impact of any student selection by collecting over-time data to account for any baseline differences between classrooms. Exploiting the as-good-as-random assignment of students to a particular classroom, I estimate the effects of being assigned to a homeroom teacher who decided to enroll in a virtual exchange program (VEP) as part of a teaching fellowship. I compare students assigned to a teacher who applies to take part in the VEP with students in the same grade, in the same school, assigned to teachers who do not apply.

Because randomization occurs at the classroom level, the treatment entails virtual contact as well as other curricular choices made by the teachers. I partly mitigate this concern by showing baseline balance on teacher attitudes that proxy for a commitment to diversity, collecting baseline student data a third of the way into the school year such that some teacher effects are absorbed, and demonstrating similar results for treatment-and-control pairs who share the same teacher. Nevertheless, treatment effects likely represent an upper bound, and teachers committed to diversity form a critical part of the intervention.

Students assigned to classrooms that participate in video exchanges with classrooms in other parts of the world are more likely to perceive commonalities with foreigners (13%,  $p < 0.09$ ), Brown children (11%,  $p < 0.16$ ), Black children (20%,  $p < 0.05$ ), and White children (26%,  $p < 0.01$ ). Contact works through social and emotional channels rather than intellectual ones: treated students are more likely report tolerant norms among their friends (10%,  $p < 0.09$ ) and more likely to practice empathy (14%,  $p < 0.10$ ), but not more likely to report knowing more about life in other parts of the world (2%,  $p < 0.81$ ). These effects are driven by two trends: control groups growing more prejudiced over time, and exchanges that bring together ethnically similar classrooms. The results suggest that being assigned to a classroom enrolled in virtual exchanges, led by a teacher committed to diversity, can insulate children from rising prejudice in their communities, and that ethnic overlap moderates the strength of this relationship.

## **Online Contact to Offline Tolerance?**

Like adults, children are prone to prejudice and in-group bias. Children start to favor their own social group as soon as the basic processes of social categorization and identification emerge in early childhood, peaking among 5

– 7 year olds (Levy and Killen 2008; Raabe and Beelmann 2011; Baron and Banaji 2006; Aboud 2003; Beelmann and Heinemann 2014; Bigler and Liben 1993; Doyle and Aboud 1995). Susceptibility to social norms partly explains this relationship, with younger children more likely to learn and imitate prejudice from parents, peers, and teachers (Vezzali, Giovannini and Capozza 2012; Aboud and Doyle 1996; Epstein and Komorita 1966). Once formed, prejudice tends to stagnate throughout adolescence (Beelmann and Heinemann 2014).

Can virtual contact reduce prejudice among children? Several meta-analyses converge on the potential for direct intergroup contact to mitigate prejudice by highlighting commonalities, filling information gaps, and inducing empathy (Aboud et al. 2012; Pettigrew and Tropp 2006; Paluck, Green and Green 2017). The conditions needed to activate the contact hypothesis, however, remain unclear. Contact is thought to be effective when endorsed by communal authorities and norms, places participants on equal footing, and requires cooperation toward a common goal (Allport, Clark and Pettigrew 1954), but the importance of these conditions, as well as the importance of physical proximity, remain unknown.

Taking contact online might be futile for two reasons. First, physical proximity is important for cooperation, empathy, and trust. Experimental evidence shows that physical touch increases cooperation, suggesting that in-person contact can build trust in ways that virtual contact cannot (Schroeder et al. 2014). Findings from neurobiology indicate that in-person contact is needed to activate mirror-neurons, which form the building blocks of imitation and empathy in humans (Rizzolatti and Craighero 2004). In-person contact is also more conducive to picking up on non-verbal cues and thus more communicative (Mehrabian 2017). Second, research on contact among children concludes that contact alone is insufficient. Success critically depends on whether contact programs initiate intergroup friendships (Pettigrew 1997; Aboud and Levy 2000; Christ et al. 2010). Meta-analyses of studies among children (Davies et al. 2011) confirm that these personal relationships are “impossible or at least difficult to promote in an indirect contact situation,” such as virtual contact (Tropp and Prenovost 2008).

Conversely, virtual contact, like its traditional counterpart, could reduce prejudice through three pathways based on information, empathy, and social norms (Scacco and Warren 2018). First, virtual contact can fill knowledge gaps where information about outgroups was inaccurate, deficient, or non-existent. Contact can convey factual information about life in other communities, which has been shown to make children more tolerant (Nasie and Diesendruck 2019). Second, virtual interactions may not be amenable to building personal friendships, but they can nonetheless induce empathy by humanizing an outgroup and showing them to be full, complex individu-

als relative to the simplified heuristics that characterize prejudice (Sniderman 2015). For instance, students in the study sample from Lévis, Québec, one of the most ethnically homogenous cities in Canada with a 99.98% European-descent population (Statistics Canada 2017), can glean from their exchanges with other fourth-graders in New York City that foreigners are “are lots of fun and are a lot like me,” despite 83% of this class reporting knowing “nothing” or “little” about life in other countries. A blurry image of children from other parts of the world comes into focus, on both the cognitive and emotional planes.

Third, virtual contact can promote tolerant social norms. Our perceptions of whether those around us approve or disapprove of intergroup contact is a “strong motivating force” underlying intergroup attitudes and behaviors (Pettigrew and Tropp 2013, p. 92), even in indirect contact settings (Paluck 2009; Pettigrew et al. 2007; Wright et al. 1997). Observing friends and teachers engaging positively with outgroups is a credible signal that tolerant attitudes and behavior are acceptable or even encouraged (Turner et al. 2008; Pettigrew and Tropp 2013; Tropp et al. 2016). Although overt racism is increasingly rare (Sniderman and Tetlock 1986), children’s implicit prejudice nonetheless correlates strongly with the implicit prejudice of their teachers, suggesting that children pick up on unspoken intergroup norms (Vezzali, Giovannini and Capozza 2012). For these reasons (among others), the contact hypothesis holds that supportive norms are necessary for contact to reduce prejudice (Allport, Clark and Pettigrew 1954). Since parents are not part of the VEP, but classmates are, I should observe effects on how students perceive norms among their classmates, but not among their parents.

Can virtual contact reduce ethnic prejudice in particular, as opposed to prejudice more generally? Interventions aimed at reducing ethnic or racial prejudice generate “substantially weaker effects” relative to prejudice toward other groups like the elderly or the disabled (Paluck, Green and Green 2017). Children are also naturally inclined to fixate on ethnicity. As soon as children start to distinguish between ethnic categories, they begin to perceive outgroup members as different, infer that they have different preferences and ways of life, and subsequently judge them as ‘wrong’ (Killen and Rutland 2011). They also respond strongly to norms set by the adults around them (Vezzali, Giovannini and Capozza 2012; Skinner, Meltzoff and Olson 2017). The sensitivity of children to social norms cuts both ways: teachers in the study sample self-select into the VEP based on their commitment to diversity, but intolerance has been rising in their broader communities. About two-thirds of classrooms in the study sample are located in counties that voted for President Trump in 2016, a watershed moment in U.S. politics that normalized prejudice toward ethnic minorities and foreigners (Schaffner 2018). For many students in the sample, despite the

best efforts of their teachers, the social and political environment leans toward prejudice when they step outside the classroom.

Given these challenges, I propose an important limitation to the impacts of contact. The ability of contact to reduce prejudice toward an outgroup writ large — rather than the specific individual one has come into contact with — cannot be taken for granted. One can easily build a mental fence around an individual classmate, friend, or acquaintance from an ethnic outgroup, and view her as exceptional. We know that this inference process is mediated by political factors like hateful rhetoric from political parties and national leaders (Sønderskov and Thomsen 2015; Hopkins 2010; Ferwerda, Flynn and Horiuchi 2017). This aligns with the finding that exposure to conflict solidifies group identities, ethnic prejudices, and anxieties around being physically proximate to the outgroup (Scacco and Warren 2018; Fearon and Laitin 2000; Beber, Roessler and Scacco 2014; Enos and Gidron 2016; Kunovich and Hodson 2002). We also know, however, that direct personal experience has the power to override the negative influence of environmental factors in shaping opinions, at least for political attitudes (Lerman and McCabe 2017; Lerman and Weaver 2014; Mettler and Milstein 2007; Soss 2002; Campbell 2003). Can personal experience do the same for social attitudes?

Tying these insights together, I argue that the ability to update positively on an entire group based on personal experience with people from that group is driven, in part, by the salience of group conflict at baseline. The more salient the conflict, the riskier it is to trust or empathize with strangers from the outgroup, who may treat my ingroup negatively. When intergroup tensions are high, the expected value of any given interaction is negative. Stereotypes, skepticism, and ingrained social norms will determine how one treats strangers from the outgroup, rather than direct experiences, even if those experiences are positive. The less salient the conflict, however, the safer it is to empathize with outgroup strangers. When ethnicity is not top-of-mind, information provided by direct experiences will overwhelm any prejudicial messaging disseminated by opinion leaders. This implies that prejudice reduction interventions are least effective where they are needed the most, but are still potentially more effective than no intervention at all.

I decompose conflict salience itself into two parts. First, the perceived riskiness of dealing with the outgroup members could be a function of whether or not one belongs to a stigmatized minority. For instance, Black children are more concerned about being being the target of bias in interracial interactions, which dilutes their enthusiasm for imagined future contact, relative to their White peers (Doerr et al. 2011). Especially when groups are highly

unequal in power, intergroup contact can devolve into frustration and disappointment for minorities (Saguy, Tropp and Hawi 2012). This lop-sided riskiness of empathizing with the outgroup may explain why contact programs are most effective for majority group children (Beelmann and Heinemann 2014), who are also quicker to point out commonalities relative to members of disadvantaged groups, who are more cynical during intergroup interactions (Saguy, Dovidio and Pratto 2008). Second, given the heightened stigmatization of ethnic outgroups and foreigners in the 2016—2019 political moment, contact with these groups in particular may increase conflict salience. The less the ethnic and national overlap with an outgroup, the more salient the conflict, and in turn, the more that contact may highlight differences rather than commonalities. These differences make it easier to revert back to the skeptical norms, stereotypes, and messaging in our environments when forming intergroup attitudes.

In sum, I posit that virtual contact, like direct contact, can reduce prejudice by increasing information, inducing empathy, and fostering tolerant social norms. These potential returns, however, are limited by the salience of intergroup conflict. The more salient are ethnic differences, the less likely it is that contact can change generalized attitudes about the outgroup as a whole. This conceptual framework gives rise to several observable implications. If contact increases information and empathy, then it should move outcomes that capture knowledge of other cultures and cognitive empathy, respectively. If contact builds tolerant norms, then outcomes that capture students' perceptions of how their friends would react to their befriending of a foreigner should increase as a result of the treatment, whereas norms about how their parents — who are not part of the VEP — would react should be less affected.

Finally, if the positive returns to contact are mediated by conflict salience, then we should expect the strongest effects among majority-group students, for whom intergroup tensions and ethnic identity are less top-of-mind, and the weakest effects among minority students, who have more to lose by empathizing with higher-status outgroups. Likewise, we should expect the weakest effects for exchanges that bring together socially distant classrooms with little ethnic overlap, where intergroup differences are brought to the fore. Alternatively, conflict salience might play a negligible role in shaping the effects of contact. If the positive social norms set by the teacher in addition to children's innate curiosity win out, then we should expect no differential effects based on a students' minority-majority status, or based on the social distance of the virtual match.

## **Study Setting**

### **Virtual Exchanges**

To activate the prejudice-reducing potential of intergroup contact, such contact should involve cooperating for a common goal, an equal power status between the two groups within the context of the intervention, and the “support of authorities, law, or custom” (Allport, Clark and Pettigrew 1954). The VEP evaluated here was designed to foster intergroup contact that meets these conditions: students cooperate to work through activities and achieve learning goals set by the teacher, exchanges are endorsed by authority figures (teachers) on both sides, and students share an equal power status within the exchange. The main source of difference between this VEP and the original contact hypothesis is therefore the medium of video-based engagement.

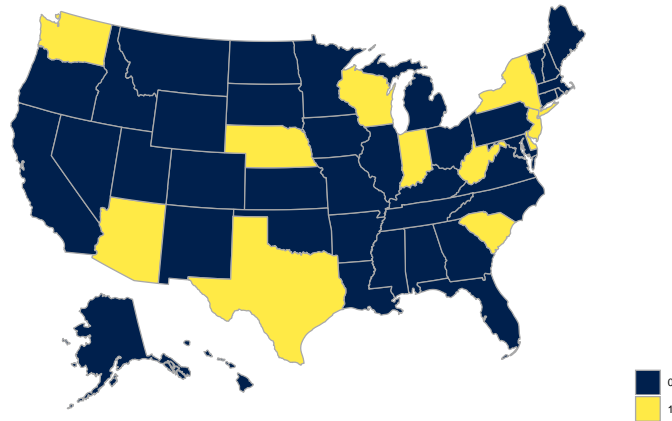
The Empatico online platform ([www.empatico.org](http://www.empatico.org)) is a free video exchange tool aimed at improving knowledge, empathy, and understanding of different cultures between children aged 6 – 11. The platform consists of teacher resources, student activities, and a series of video exchanges with classrooms in other parts of the world. The platform matches classrooms based on two criteria: scheduling availability, and preference for one of twelve educational activities for the paired classrooms to work through together. Teachers typically choose a given activity based on their class’s curricular needs. For instance, a teacher might choose the ‘community cartographers’ activity to fit a geography class, or a ‘festivals around the world’ activity to suit a social studies class. Based on these two constraints — which are unrelated to tolerance – the platform creates a match. Once a match is made, teachers conduct exchanges using the platform’s native video conferencing tool.

### **Classroom Recruitment**

Several teaching fellowships incorporate Empatico into their fellowship program. The classrooms I evaluate are led by elementary school teachers who hold a fellowship at Empatico itself (46.6%) or through the National Network of State Teachers of the Year (NNSTOY) program (53.4%). Teaching fellows received a stipend of \$750, and, in return, conduct at least four exchanges with partner classrooms, in addition to regular de-briefing sessions with other fellows. The high degree of organization, structure, and incentives embedded in the fellowship ensured perfect compliance and high retention rates, with all teaching fellows retained until the end of the study. This selection process resulted in a modal student who goes to a Title 1 school (53.6%) located in U.S. (85.0%), and



**Figure 1:** States Represented in U.S. Study Sample



All states have one treatment-control pair except Wisconsin, which has two pairs.

receives a free or reduced price lunch (55.3%). The modal classroom is ethnically homogenous (80% of students or more sharing one ethnicity), with roughly 70% of students in majority white classrooms, 10% in majority Black classrooms, and 20% in majority Hispanic classrooms. Of the American classrooms, 11 states are represented (Figure 1). Of the international classrooms, two experimental pairs are located in Mexico, and one pair apiece is located in Canada, Honduras, and Israel.<sup>1</sup>

## Empirical Strategy

### Experimental Design

What are the *causal* impacts of virtual contact? Two features of the study setting and research design allow me to construct the plausible counterfactual needed to answer this question. First, the assignment of most students to a particular homeroom teacher at the start of the year is as-good-as-random. Students are not able to choose their own teachers, and parents are unaware of the teacher’s plans to enroll in the VEP until after the school year begins. Schools may, however, make decisions over classroom assignment based on specialized learning needs for some students, namely students of accelerated learning ability or students that require special support. 95.2% of students in the sample are described by their teachers as students of “traditional learning ability,” suggesting

---

<sup>1</sup>Although based in Israel, these two classrooms are American international schools serving primarily American students.

that interference in class assignments to accommodate student abilities is minimal. Second, even if there is some selection of students into particular classrooms, perhaps driven by demand effects for certain teachers, over-time data mitigates concerns about any remaining baseline differences between these two groups of students.

I construct a control group in two ways. First, some teachers teach multiple classes of students at once. These multi-class teachers offer an ethical opportunity to withhold the treatment from one of these classes, as the Empatico platform only allows teachers to enroll one class at a time. Importantly, multi-class teachers select a particular class to participate in the program based on scheduling availability or curricular needs, rather than on the basis of traits like empathy or knowledge of other cultures, which correlate with intergroup attitudes. 24 of the 27 baseline surveys of multi-class teachers reveal that teachers choose a particular class on one of three bases: scheduling availability, a set of students fitting the target age-range, and the classroom that they spend most time with (e.g. homeroom).<sup>2</sup> By comparing students taught by the same teacher (15% of the student sample), I hold constant teacher traits like personality, experience, and teaching style. Any differences observed between students can then be attributed to the VEP. For teachers who teach one classroom, I create a control group using a different method. Single-class teachers nominate a “companion class” to participate in the study. A companion class is another class in the same grade, in the same school, that does not plan to use the VEP during the study period. This recruitment strategy achieved balance on key classroom and teacher level traits (Figure 2).

All of the 66 teaching fellows completed a baseline and endline survey, and invited a control classroom to participate in the study. 24 control classrooms went as far as obtaining parental consent and administering the baseline survey. Of these, four classrooms were not able to complete the endline surveys before the summer break began.<sup>3</sup> The final sample thus consists of 20 complete treatment and control pairs who were also able to complete baseline and endline surveys, yielding a total of 42 classes (30.3% of invited classes).<sup>4</sup> I keep only complete treatment-and-control pairs, who completed both the baseline and endline surveys. This approach accounts for any baseline differences between the treatment and control classrooms, ensures that the composition of treatment and control groups before and after the intervention is stable (an assumption for the difference-in-differences strategy), and increases precision.

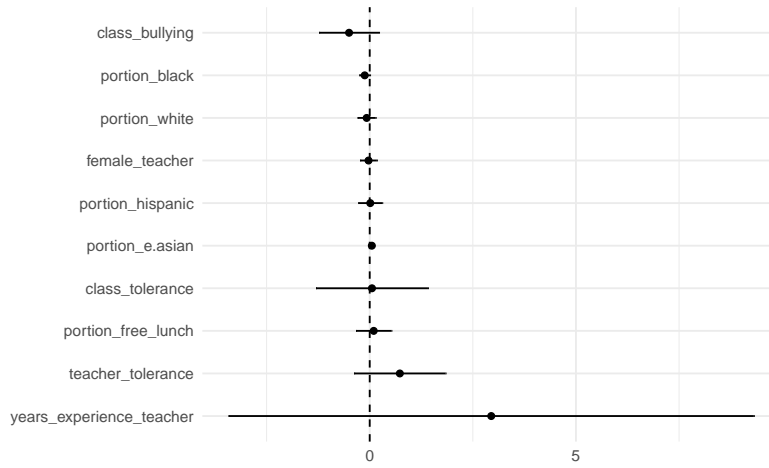
---

<sup>2</sup>The remaining three teachers reported that learning about other cultures aligns with the class’s social studies curricula or learning progression.

<sup>3</sup>Endline surveys were taken two weeks after the treated classroom’s final exchange, which coincided with the hectic final week of classes for many schools.

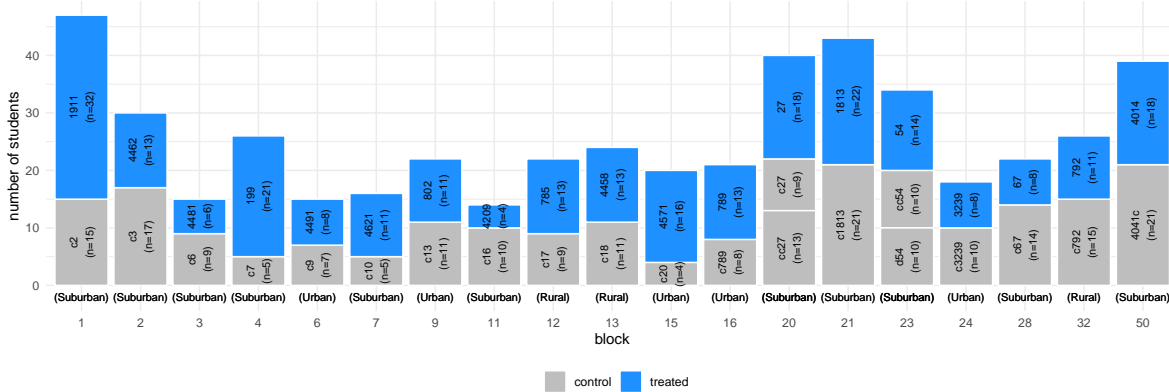
<sup>4</sup>Two treated class enrolled two control classes, see Figure 3.

**Figure 2: Balance Across Treatment and Control Classrooms ( $n = 42$ )**



Measures taken from baseline teacher surveys, and reflect teacher traits (gender, tolerance, and years of experience) and classroom traits (portion of Black, White, Hispanic, and East Asian students, portion of students availing free or reduced price lunch, and the prevalence of two classroom norms: anti-bullying, and tolerance of ethnic and socio-economic difference).

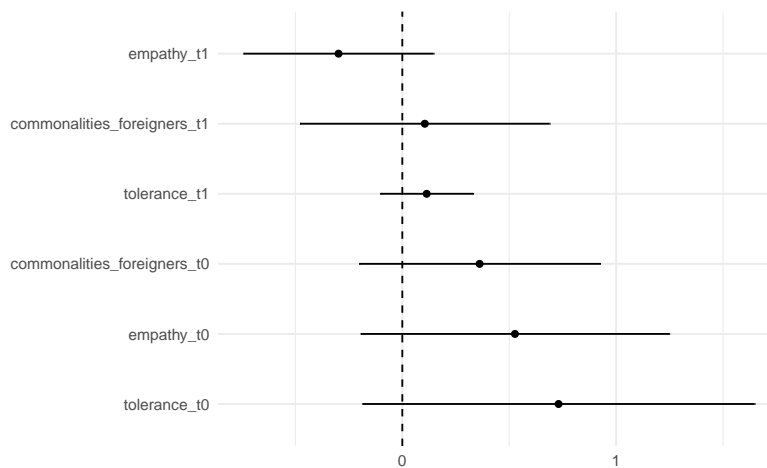
**Figure 3: Study Sample by School Location ( $n = 508$ )**



This empirical strategy causally identifies the effects of the treatment — but this treatment is bundled. Teachers who sign up for virtual exchanges have revealed their preference for instilling the values of tolerance, diversity, and empathy in their students. Teachers who enroll their classes in virtual exchanges plausibly make other efforts to incorporate these values into their classrooms. Three sources of data speak to this concern. First, all but one classroom took the baseline survey between mid-November and mid-December. Most students were therefore two to three months into their classroom assignments (25 – 30% as a proportion of the school year) before baseline

data is collected. At least some of the prejudice-reducing efforts individual teachers embarked upon are therefore absorbed in the baseline data. Second, Figure 4 shows that treatment and control teachers share similar views on empathy, diversity, and tolerance at baseline and at endline, an ideological similarity underscored by control teacher who noted: “I was asked by a professional peer to participate [in the survey] and agreed because I agree with the ideals of this organization.” Third, I analyze data among multi-class teachers, which make up 6 of the 42 study classrooms. Both the treatment and control group have the same teacher for this subset of the sample. Some imprecision in the second and third analyses is attributable to the small sample size, but these analyses nevertheless rule out that there are systematic, large differences between classrooms by treatment status.<sup>5</sup> While these pieces of evidence suggest that the program effects do not rest entirely on teacher choices outside of virtual contact, generalizability is likely contingent on virtual contact led by motivated, diversity-minded teachers.

**Figure 4: Teacher Balance on Intergroup Attitudes ( $n = 34$ )**



Estimates based on a two-sided t-test. Outcomes are binary and coded in a tolerant direction, and reflect agreement with the following statements: “I believe that there are two sides to every situation, and I always try to look at them both” (empathy), “This country would be a better society if people treated each other as equals first, rather than divide people by citizenship, religion, or ethnicity” (tolerance), and “I share a lot in common with people in other countries” (perceived commonalities with foreigners). Error bars represent 90% confidence intervals.

How do ethnic differences between the matched classrooms shape the impacts of contact? I leverage a natural experiment built-in to the Empatico platform to shed light on this question. The Empatico matching algorithm automatically finds a partner class based on scheduling availability and topics of interest (e.g. the local environment, or holidays around the world). Conditional on factors that are orthogonal to tolerance, social distance to

<sup>5</sup>Minimum detectable effect calculations to follow.

the partner classroom is therefore randomly assigned. This measure is binned into two (pre-registered) levels: (1) fewer than 25% co-ethnic, and (2) over 25% co-ethnic. Pairs that bring together classrooms in two different countries are categorized as non co-ethnics. This information is taken from the teacher surveys, although I use district-level ethnic composition information for four classrooms where no teacher surveys were available.<sup>6</sup>

## Estimation

I estimate the average treatment effect (ATE) on a range of attitudinal items in three ways. First, for the main analysis, I analyze the ATE using a difference-in-differences (DID) specification. The DID estimator subtracts the baseline levels from the endline, and preserves all outcome variables as continuous. Second, I use an ordinary least squares (OLS) estimator where the treatment indicator represents assignment to a take part in video exchanges, which occurs at the classroom level. Outcomes are binary for this analysis, and the cut-off for each outcome represents moving from a negative or neutral response to an explicitly positive response. Third, to deal with estimates that exceed the boundary of 1, I run a logistic regression. I present the second and third specifications as separate analyses presented in the appendix.

For all three approaches, I take several demographic items as covariates to increase precision, in addition the outcome variable measured at baseline where possible, and the randomization block. The demographic covariates are: country, language spoken, whether the student has siblings (known to correlate positively with empathy (Lam, Solmeyer and McHale 2012), whether they have conducted video exchanges before, and ethnicity, proxied by the student's self-reported skin tone, asked last to avoid priming effects (Figure 14). Surveys are administered twice: two weeks before the first exchange, and two weeks following the final exchange of the school year. I cluster standard errors at the classroom level.

I use subgroup analyses to interrogate the importance of conflict salience. I do so by analyzing heterogeneous treatment effects based on three dimensions measured at baseline. First, whether the exchange involved a 'socially distant' partner classroom, where fewer than 25% of students share an ethnicity or nationality. I take ethnicity as the most obvious source of perceived social distance among children, who may not be able to pick up on more subtle sources of difference, such as socio-economic status. Ethnicity is socially constructed, but humans instinctively process ethnic differences as if they were essential Gil-White (2001). I gain causal leverage

---

<sup>6</sup>These four schools are located in: Lafayette, IN, Elgin, SC, Camden, DE, and Madison, WI.

on this question by exploiting the VEP's matching process. Conditional on scheduling availability and curricular preference, which are unrelated to tolerance, the VEP randomly pairs two classrooms together to engage in exchanges. To test the importance of social distance, I compare the trajectories of students randomly paired with a predominantly non-coethnic classroom, to those paired with students who largely look like them.

Second, I break down results by whether the student belongs to the classroom's ethnic majority or not, which captures majority status at the classroom level. Third, I do the same based on whether the student self-identifies as white, reflecting majority status at the national level. I interact these variables with the treatment indicator, and include the lower order term, for these subgroup analyses. These outcomes, along with baseline covariates, and heterogenous treatment effects, were pre-registered prior to data collection with Evidence in Governance and Politics (EGAP registration #20190215AA). Finally, missing data for baseline covariates (but not outcomes) are imputed using multiple imputation by chained equations. 6.82% of rows contain a missing value requiring imputation.

## **Data and Outcome Measures**

I measure a range of attitudinal and behavioral outcomes among students, using both teacher and student surveys to triangulate information. Starting with attitudes, the student survey includes validated items that measure cognitive empathy (perspective-taking), self-other overlap (perceived similarity with a specific out-group, (Gómez, Tropp and Fernández 2011), perceived acceptability of having an outgroup friend among one's friends and parents (social norms), and knowledge of life in other countries (Alvídrez et al. 2015; Cameron et al. 2001; Rowley et al. 2008; Tropp, O'Brien and Migacheva 2014; Beelmann and Heinemann 2014). A meta-analysis of 211 contact interventions among children concludes that choosing to spend time with out-group members (e.g. via offline communication) is a valid behavioral measure of attitudes toward social outgroups (Davies et al. 2011), as is indicating sustained interest in communicating (West et al. 2014). The survey instrument therefore includes an item on interest in having a foreign penpal in the future, and interest in meeting children from other countries. These survey-based measures are valid but not always reliable among young children (West 2016). This measurement error adds noise to the estimation, but not in a way that should systematically differ by treatment status.

I also measure three student behaviors. First, I run a sentiment analysis on open-ended questions asking students what they think of people from other countries. I measure the portion of positive and negative words used to

describe foreigners using the three lexicons embedded in the `tidytext` package: Bing, AFINN, and NRC. The Bing lexicon classifies words as positive or negative, the AFINN lexicon assigns words a sentiment score from  $-5$  to  $5$ , and the NRC lexicon categorizes words into one of ten sentiment buckets (positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise and trust). I collapse these classifications into binary categories that represent positive or negative sentiments.<sup>7</sup> The weakness of relying on such lexicons is that sentiment is often context dependent. For instance, one student wrote that ‘they [kids from other countries] are just like us but with a twist.’ Despite the overall positivity of the statement, the word ‘twist’ is categorized as negative by the lexicons. On the other hand, such lexicons provide an objective, albeit noisy, indication of sentiment. This approach is arguably preferable to hand-coded sentiment scores, as it would be nearly impossible for a coder to be totally blind to treatment status given that many treated students refer to their partner classrooms in their responses.

Second, the teacher survey includes three items capturing student behavior: the degree to which bullying is a problem, the degree to which students stop others from being unfair or disruptive, and the degree to which students get along with students of other social classes and races. These items are measured at the classroom rather than individual level. Third, among students that take part in exchanges, I measure whether they write or draw a message to send to peers in their partner classroom. I do this through teacher-delivered prompts at the end of the semester, one week after the final exchange: “if you would like to send a message to a student in our partner classroom, please write it out on a piece of paper, and I will take a photo of it to send to them.” I track which students choose to write out and send such a message. This outcome is measured only among the treated group, but grants me the opportunity to explore how individual-level factors, such as ethnicity, as well as classroom-level factors, like the identity of the partner classroom, correlate with the writing of penpal messages.

## Results

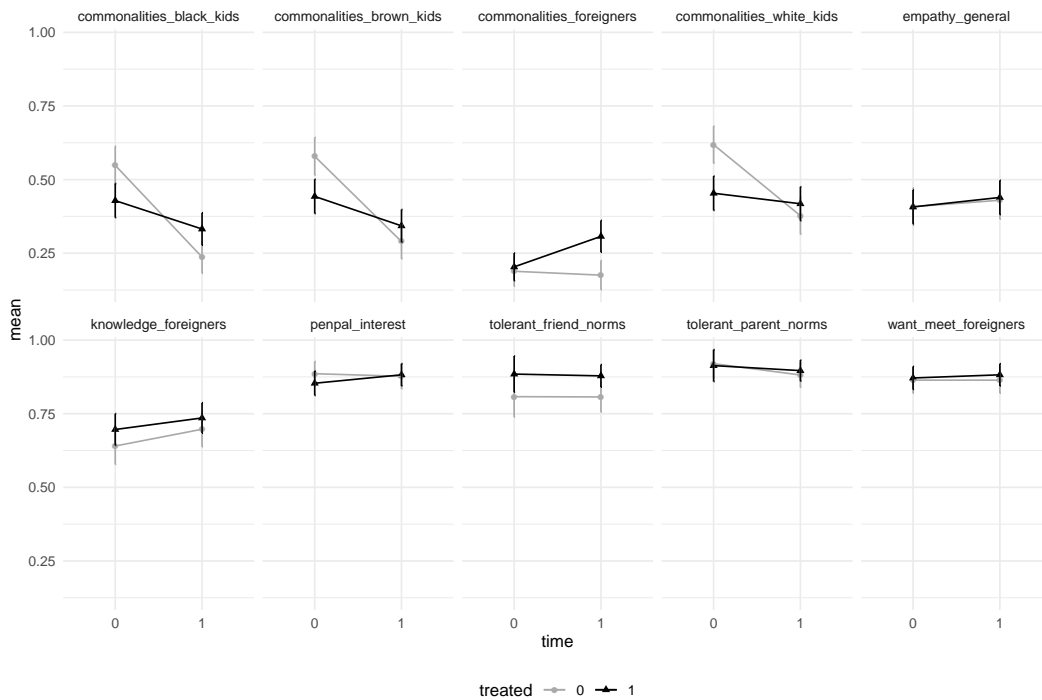
Figure 5 plots the raw means for each treatment group and time period, Figure 6 visualizes the main results, and Table 1 contains the OLS-adjusted results. Starting with intergroup attitudes, assignment to treated classrooms made students more likely to report commonalities toward all four outgroups asked about in the survey: Brown children (13%,  $p < 0.09$ ), Black children (19%,  $p < 0.06$ ), White children (22%,  $p < 0.02$ ), and foreign children (11%,  $p < 0.09$ ). These improvements equate to a shift from a negative or neutral response to an explicitly positive

---

<sup>7</sup>I classify the positive, anticipation, joy, surprise, and trust categories as positive, and negative, anger, disgust, fear, and sadness categories as negative.

response for Black children and foreigners, who were the target of the most prejudicial attitudes at baseline (Figure 5), as well as an increase in tolerant norms and empathy (Figure 10. Table 2), Treatment effects on interest in having a foreign penpal and in befriending foreigners are positive but not robust across the model specifications (Table 2, Figure 6, Figure 10).

**Figure 5: Raw Means by Treatment Condition**

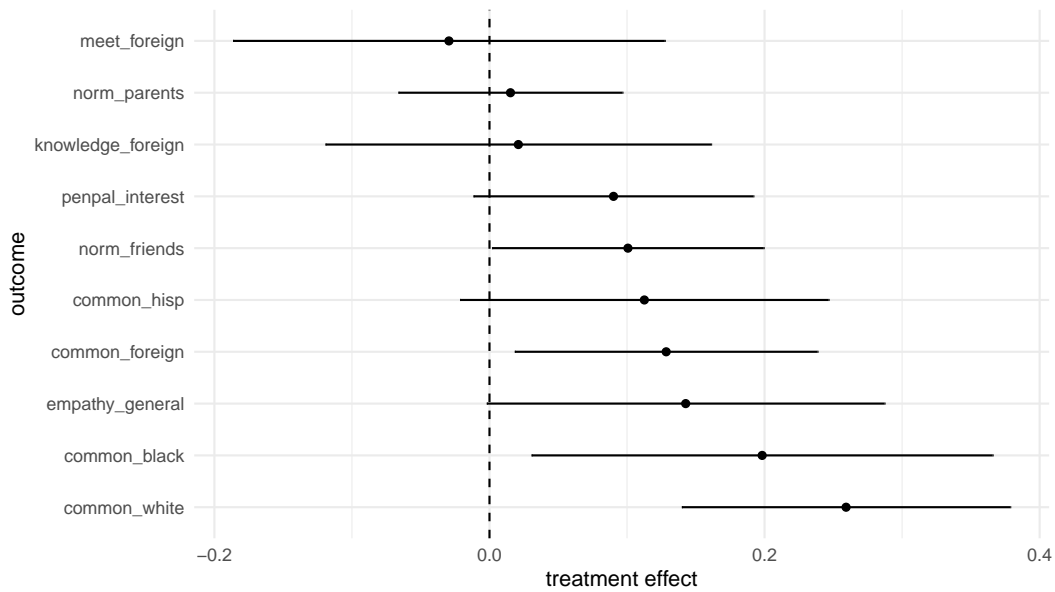


Outcomes are binary, and can be interpreted as the proportion of students in each condition with a positive attitude. Neutral responses are coded as a 0.

Moving to behaviors, virtual exchanges increased the portion of positive words used to describe children from other countries for two of the three lexicons used (Figure 7) at an effect size of 30 – 35 percentage-points. Looking at raw means, the share of positive words at endline ranges from 16.6 — 23.4% for the control group and 23.8 — 27.7% for the treated group, depending on the lexicon. Virtual exchanges also increased the portion of negative words used to describe this outgroup, but this effect is negligible; the average share of negative words used at endline ranges from 0.02 — 0.03% for both the treated and control group. The final set of outcomes reflects classroom-level behaviors recorded by teachers. The teacher sample is too small to draw reliable conclusions ( $n = 87$ ), but teachers of treated classroom teachers report a roughly 30% improvement to classroom norms around



**Figure 6:** ATE: Difference-in-Differences ( $n = 508$ )



Treatment coefficients based on OLS regressions where outcomes are continuous, and differenced-out ( $t_1 - t_0$ ). Models include the following covariates: randomization block, outcome measured at baseline, language spoken, country, if one has a sibling, and ethnicity (proxied by self-reported skin tone). Error bars represent 90% confidence intervals.

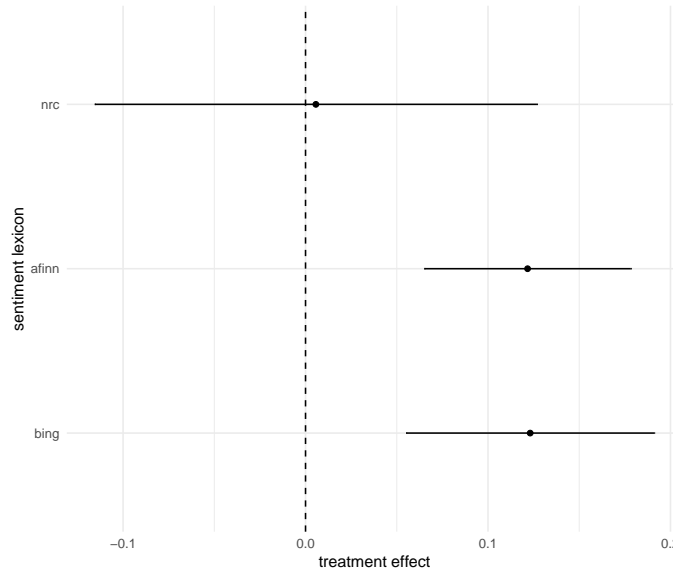
bullying ( $p < 0.01$ ) and tolerating difference ( $p < 0.07$ ), and a 10% reduction in bullying itself, although this change is not statistically significant ( $p < 0.19$ , Figure 13).<sup>8</sup>

The results affirm the idea that contact works through social and emotional channels. Students who took part in the virtual exchanges were marginally more empathetic (11 %,  $p < 0.15$ ), as measured by an item on perspective-taking, and more likely to report that their friends would accept an intergroup friendship (13%,  $p < 0.06$ ), but not their parents, as expected for a classroom intervention (1%,  $p < 0.82$ ). In contrast, the relationship between contact and increasing information about life in other countries is positive but not robust across specifications (Figure 6, Figure 10). Five of the six attitudinal results survive the Benjamini-Hochberg multiple comparisons correction at the  $\alpha = 0.1$  level (Benjamini and Hochberg 1995).<sup>9</sup>

<sup>8</sup>These questions are worded as follows, and measured on a 5-point Likert scale: (1) “How much do you agree with the following statement: ‘Students in my classroom stop other students who are unfair or disruptive’”; (2) “How much do you agree with the following statement: ‘Students from different social classes and races get along well;’” and (3) “How much do you agree with the following statement: ‘Bullying is a problem in my classroom.’”

<sup>9</sup>The treatment effect on empathy ( $p < 0.15$ ) does not survive this correction. When setting  $\alpha = 0.05$  level, four of the six attitudinal results remain, with the result on commonalities with foreigners  $p < 0.09$  failing to meet the critical value of 0.05.

**Figure 7: Sentiment Analysis**



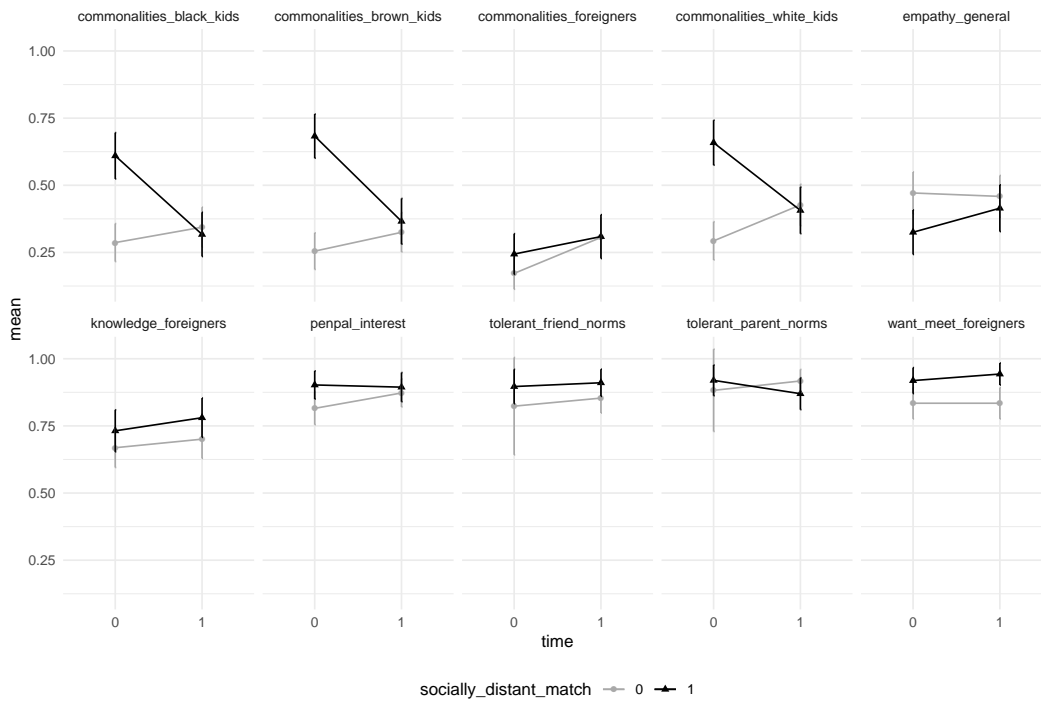
Outcomes range from 0 to 1 and reflect the portion of positive words written by each student when asked what they think about people from other parts of the world. Estimates are based on an OLS model controlling for treatment block, classroom, student ethnicity, language spoken, having siblings, and country. Standard errors are clustered at the student level. Error bars represent 90% confidence intervals.

These results are driven by two factors in particular. First, the raw means reveal that the attitudes of children in control classrooms grew more negative over time for seven of the ten attitudinal outcomes (Figure 5). The attitudes of children in treated classrooms, by contrast, remained stable or become more positive. Second, among the treated students, positive results are concentrated among those who connected with socially proximate classrooms, whereas connecting with socially distant classrooms — meaning a  $\leq 25\%$  overlap in ethnicities or nationalities — diluted attitudinal changes relative to contact between socially similar classrooms (Figure 8). The added effect of being assigned to participate in video exchanges with a socially distant partner is negative for all ten outcomes, and this negative interaction is statistically significant for five of the ten outcomes, where the magnitude of this effect ranges from around  $-7 - -20\%$ . (Table 4). Virtual contact with a socially distant classroom was still better, however, than no contact at all (Figure 9).

I find comparatively weaker evidence for the hypothesis that minority group status, the other theorized dimension of conflict salience, matters for mediating the effects of contact. Identifying as white or belonging to a white-majority classroom is associated with more tolerant attitudes before the intervention, and amplifies treatment

effects for some outcomes, but not more than what we would expect due to chance (Tables 6 and 7). The reverse is true for ethnic minorities, who hold generally less tolerant attitudes at baseline, and for whom treatment effects are weaker when interacted with their minority status (Table 5). White students were also more likely to write a message to a foreign penpal relative to their non-white counterparts (13% vs. 9%), but this difference in means is not statistically significant ( $p < 0.30$ ). The relationship between minority status and contact effects moves in the predicted direction, but is not systematic enough to draw an inference.

**Figure 8:** Raw Means by Social Distance (treatment group only;  $n = 280$ )



Sample restricted to the treated group, and divided by students who had a socially distant (less than 25% ethnic overlap) vs. socially proximate (25 — 100% ethnic overlap).

## Discussion

The results show that, on average, classrooms led by teachers who choose to partake in virtual exchanges can stave off rises in prejudice taking place outside of the class walls, primarily by highlighting commonalities, inducing empathy, and strengthening tolerant social norms. These effects range in magnitude from 10% – 25% for intergroup attitudes and sentiment, and, although direct comparisons are difficult, are similar in scale with

studies of face-to-face intergroup contact, at an average effect size of 20% – 30% in [Pettigrew and Tropp's 2006](#) meta-analysis, and 20% among experimental tests of contact (excluding extreme results from studies targeting the disabled) in [Paluck, Green and Green's 2017](#) meta-analysis. In line with a large body of literature, the results affirm that outgroup affect is much more important than outgroup information as a mechanism mediating the effect of contact on prejudice ([Pettigrew and Tropp 2013](#)). Further, the robust treatment effect on the perceived acceptability of befriending someone from another country aligns with a rich evidence base showing that the desire to befriend those from outgroups is highly sensitive to social norms, and that observing our friends engage in intergroup contact is a strong signal of these norms ([Pettigrew et al. 2007](#); [Wright et al. 1997](#); [Dovidio, Eller and Hewstone 2011](#)). Even in settings of high intergroup conflict, such as between British-Indians and white Brits ([Turner et al. 2008](#)), Black and white Americans ([Pettigrew and Tropp 2013](#)), and indigenous and non-indigenous Chileans ([Tropp et al. 2016](#)), the repeated social rewards for engaging in intergroup contact can work to diminish prejudice. I find the same pattern here.

These gains were strongest when contact brought together classrooms with some ethnic or national overlap. The small sample size runs the risk of over-estimating the magnitude of subgroup effects ([Gelman and Carlin 2014](#)), but the systematically larger effects among ethnically similar matches in both the raw and parametric results tell a consistent story. In contrast with work showing weak or even negative effects of contact among minority group members ([Saguy, Dovidio and Pratto 2008](#); [Doerr et al. 2011](#); [Saguy, Tropp and Hawi 2012](#)), I find mixed evidence on the response of minority children to contact. These children experience somewhat negative treatment effects, but this pattern cannot be separated from noise ([Table 5](#)). Moreover, when subsetting the analysis to Black children, I find the biggest improvements to perceived commonalities with White children ([Figure 12](#)). Social distance, regardless of one's own group membership, thus seems more influential in determining the effectiveness of prejudice reduction programs.

Self-reported attitudes are noisy among young children ([West et al. 2014](#)), but not in a way that should differ by treatment status. The moderating role of ethnic overlap thus warrants some interpretation. A theory that favors conflict salience sheds some light on this finding. Prior research demonstrates that the political context can influence the outcome of intergroup encounters, and shape the process of generalizing from specific encounters to broader attitudes and behavior ([Sønderskov and Thomsen 2015](#); [Hopkins 2010](#); [Ferwerda, Flynn and Horiuchi 2017](#)). In this vein, I propose that contact generally chips away at prejudice, but that this effect is driven by contact with outgroups who different enough to be interesting and enlightening, but not so distant as to activate

prejudicial heuristics ingrained by our socio-political environments. While [Sønderskov and Thomsen \(2015\)](#) find that a hostile political climate can enhance the effects of intergroup contact, the present study lacks the intergroup friendships and workplace connections that the authors and others ([Wagner et al. 2003](#); [Andersson, Dehdari et al. 2020](#)) deem necessary to unlock these effects. Even with limited friendship potential, however, treatment effects were still positive on the whole, in keeping previous research on the benefits of contact that highlights both differences and similarities ([Ioannou, Hewstone and Al Ramiah 2017](#)).

The increase in prejudice among the control group was an ex-post discovery. With this in mind, conflict salience potentially points to one possible reason for this finding, echoing a similar experimental result in Nigerian classrooms ([Scacco and Warren 2018](#)). The study period (October 2018 to Spring 2019) coincided with a tumultuous time for ethnic tensions in the U.S. Ten of the eleven U.S. counties in the study sample voted for President Trump in 2016, who stigmatized ethnic minorities, immigrants, and foreigners during this time. Prejudice toward these groups became normalized under President Trump, while attitudes toward groups not targeted by his campaign (e.g. alcoholics, atheists) remained unchanged ([Schaffner 2018](#)). The FBI reported that hate crimes reached a 16-year high in 2018 ([Hassan 2019](#)), while 75% of Black Americans claimed that racism had gotten worse under President Trump ([Ballard 2019](#)). Yet only 13% of Republicans believe there is substantial discrimination against Black Americans in a 2019 Pew Research Center survey conducted six months later ([Jones 2019](#)).

Teachers noted that this normalization of prejudice trickled into their classrooms. In fact, many teachers specifically applied to teaching fellowships focused on countering these trends, which they saw as especially concerning in their communities. One teacher wrote that “today, our kids absorb a message ... that promotes suspicion of those who are different, as if different is somehow nefarious.” Another in a “historically conservative town” wrote: “Clearly, we are currently living in a divisive period of time. It is our responsibility as educators to teach our children that this behavior that is being modeled by our nation’s leaders, famous faces and citizens is unacceptable.” Teachers noted upticks in prejudice in their classrooms. One observed that “More than ever before, I am hearing students making racial slurs and this is new for me and my school,” while another cited an “increase in the number of students who have reported being bullied because of their ethnicity,” while a veteran teacher of 21 years likewise remarked that “It saddens my heart to see an increasing lack of empathy in not only our students, but also in our parents,” in support of a norms-based explanation.

Normalized hostility toward ethnic minorities and foreigners, the two outgroups in this study, may be exacerbated among children with little to no exposure to these groups in their daily lives. A teacher from a “predominantly Caucasian” school described minorities in her community as “so underrepresented that students often feel invisible.” Another from a “rural school with little diversity” stressed that “the only people of color that many [students] see are the farm workers. Otherwise, they see people of color on the evening news... which can result in misunderstandings, closed minds and intolerant thoughts.” Scholars have long noted the influence that cultural factors and mass media exert on the lived experience of prejudice among children (Milner 1983). Even for the quarter of classrooms located in urban areas (Figure 3), meaningful intergroup contact cannot be assumed. A New York City-based teacher wrote: “to the vast majority of my students, their New York is confined to the few blocks of their homes and daily routines,” while another stated that “many of my scholars have never left their block, let alone been out of Brooklyn.” Groups in conflict are more likely to be residentially segregated, making intergroup interactions less likely in the first place, and possibly worsening prejudice, offering another explanation for the deterioration of attitudes in the control group (Enos and Gidron 2016; Kunovich and Hodson 2002).

If virtual exchanges embedded in a progressive classroom environment can shield students from the negative trends outside of their classrooms, as the results here suggest, how resilient are these effects when students are no longer part of this environment? This remains an open, and important, question. Prior work shows that prejudicial social norms in the broader environment, such as the school, override positive norms found in the classroom (Nipedal, Nesdale and Killen 2010) and among peers (Paluck 2011; Tropp et al. 2016). The relative importance of communal norms relative to tolerant classroom climates should give us pause, and further emphasize the value of hiring teachers committed to fostering empathetic classroom climates.

Being assigned to a classroom led by a teacher who enrolls in virtual exchanges — a bundled treatment — insulated students from the negative trends observed in the control classroom, whatever their cause may be. How generalizable is this finding? This sample of teachers is particularly committed to the ideal of a globally connected world, the value of diversity, and instilling social and emotional skills as part of teaching the “whole child.” Aside from virtual exchanges, it is plausible (and even likely) that teachers made other efforts to incorporate these values into their classrooms. Control and treated teachers are similarly tolerant at baseline, some teacher effects are absorbed in the baseline data, and I do not detect large, systematic treatment effects for students who share the same teacher. These additional analyses suggest that virtual contact plays an important role in shaping the treatment effects, although the existing data do not allow me to pinpoint the magnitude of this role relative to other teacher-led

efforts. Committed teachers are likely critical for catalyzing the effects of virtual contact. Such teachers socialize positive norms, endorse intergroup contact, and in all likelihood embed other programs into their classrooms that shape children's' intergroup attitudes in important ways. For these reasons, we would the strongest returns to contact in classrooms led by teachers with an exceptional commitment to intergroup tolerance.

Several characteristics of the target population also help generate expectations about generalizability. The age group in the sample coincides with the peak age for childhood prejudice (Levy and Killen 2008; Raabe and Beelmann 2011; Baron and Banaji 2006), 13.0% of the teaching fellows, unprompted, used the word “poverty” to describe their students' environments, and 88.2% classified their students' socioeconomic background as ‘mostly low-income,’ factors known to correlate with prejudice (Carvacho et al. 2013). The weight of these negative student traits relative to the weight of positive teacher traits, such as high motivation and commitment to diversity, however, remains unknown. Lastly, in line with work showing the disproportionately large effects of negative contact (Graf, Paolini and Rubin 2014), contact should be optimized for a positive experience. The majority of treated students (85.7%) reported that they would want to participate in another exchange in the future, indicating that results generalize best where contact is similarly positive.

## **Conclusion**

Social geography keeps groups apart in many parts of the world, posing a challenge for using intergroup contact as a tool for building social cohesion (Enos 2014). This problem is compounded when outgroups live in other cities, regions, or countries. Virtual contact has the power to leapfrog over these structural constraints, but raises a new set of challenges, including a potentially weakened capacity to induce empathy and forge friendships, the cornerstones of traditional intergroup contact. The cost-effectiveness of video-based interactions have propelled their popularity for educational, business, and social purposes (Patterson 2019), including for 30 million school-age children in the U.S. alone in response to the COVID-19 pandemic (Strauss 2020). Yet there is little causal evidence on how virtual contact shapes social cohesion.

This study provides experimental evidence on the impacts of being assigned to a classroom that engages in virtual exchanges with peers around the world. I find that being assigned to such a classroom boosted perceived similarities with ethnic outgroups and foreigners — the target outgroups who were most stigmatized at baseline — as well as the positive sentiment used to describe these groups. In line with a rich literature, contact worked to reduce

prejudice by building empathy and strengthening tolerant norms among one's friends, rather than by increasing knowledge about outgroups (Pettigrew and Tropp 2013).

The positive impacts of being assigned to treated classrooms are driven by two underlying trends. First, effects were stronger for exchanges that brought together classrooms with some ethnic overlap, and weaker for socially distant exchanges with little to no ethnic overlap, although this type of contact was better than no contact at all. Second, as with Scacco and Warren's 2018 experimental test of contact in Nigerian classrooms, I find that children in control groups grew more prejudiced over time, possibly reflecting the increasingly intolerant climate in the U.S. during the study period. Treated classrooms, by contrast, were able to stave off these effects. Both of these findings speak to the importance of contextualizing prejudice reduction interventions within the broader macro-political environment (Sønderskov and Thomsen 2015; Hopkins 2010), while stressing the urgency of understanding the robustness of these effects when students encounter the polarized world outside their classroom.

The role played by teachers committed to moulding empathetic and open-minded children is critical in activating the positive effects of contact. When it comes to virtual contact, the endorsement of teachers signaled the acceptability of intergroup contact among students in ways predicted by the contact hypothesis (Allport, Clark and Pettigrew 1954), just as observing one's classmates interact positively with outgroups during the exchange improved perceived norms on the acceptability of intergroup friendships. Aside from virtual contact, such teachers likely infuse their classrooms with tolerant values in other explicit and implicit ways throughout the school year. While secondary analyses suggest that the main findings are unlikely to be driven solely by teacher choices outside of virtual exchanges, this study agrees with calls-to-action to leverage teachers as key stakeholders in efforts to improve intergroup relations (Beelmann and Heinemann 2014; Killen, Rutland and Ruck 2011).

That classrooms engaged in virtual contact can insulate students from rising prejudice in their environments is promising. More work is needed, however, to disentangle "teacher effects" from classroom-level interventions, and to understand the resilience of these effects when students step outside the classroom, especially on behaviors. Future work should also test the effectiveness of virtual vs. in-person contact (as opposed to no contact at all), unpack the effects of dosage, and test for lasting effects among adolescents and adults, for whom prejudice is less malleable. It remains an open question, for instance, whether more intense contact, be it face-to-face or more frequent, is needed to shift prejudice toward conflict outgroups in particular. Nevertheless, the results demonstrate



that teacher-led virtual exchanges can improve empathy, perceived commonalities, and sentiment toward ethnic minorities and foreigners — even as suspicion toward these groups appears to grow outside the classroom.

## References

- About, Frances E. 2003. "The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes?" *Developmental psychology* 39(1):48.
- About, Frances E and Anna-Beth Doyle. 1996. "Parental and peer influences on children's racial attitudes." *International Journal of Intercultural Relations* 20(3-4):371–383.
- About, Frances E, Colin Tredoux, Linda R Tropp, Christia Spears Brown, Ulrike Niens, Noraini M Noor et al. 2012. "Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review." *Developmental review* 32(4):307–336.
- About, Frances E and Sheri R Levy. 2000. "Interventions to reduce prejudice and discrimination in children and adolescents."
- Adida, Claire L, Adeline Lo and Melina R Platas. 2018. "Perspective taking can promote short-term inclusionary behavior toward Syrian refugees." *Proceedings of the National Academy of Sciences* 115(38):9521–9526.
- Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public goods and ethnic divisions." *The Quarterly Journal of Economics* 114(4):1243–1284.
- Allport, Gordon Willard, Kenneth Clark and Thomas Pettigrew. 1954. *The nature of prejudice*. Addison-Wesley Reading, MA.
- Alvídrez, Salvador, Valeriano Piñeiro-Naval, María Marcos-Ramos and José Luis Rojas-Solís. 2015. "Intergroup contact in computer-mediated communication: The interplay of a stereotype-disconfirming behavior and a lasting group identity on reducing prejudiced perceptions." *Computers in Human Behavior* 52:533–540.
- Andersson, Henrik, Sirius H Dehdari et al. 2020. Workplace Contact and Support for Anti-Immigration Parties. Technical report Centre for Research and Analysis of Migration (CReAM), Department of ?
- Ballard, Jamie. 2019. "75% of Black Americans say racism in the US has gotten worse under Trump." *YouGov* .
- Barnhardt, Sharon. 2009. "Near and Dear? Evaluating the Impact of Neighbor Diversity on Inter-Religious Attitudes." *Unpublished working paper* .
- Baron, Andrew Scott and Mahzarin R Banaji. 2006. "The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood." *Psychological science* 17(1):53–58.
- Bayer, Patrick, Robert McMillan and Kim Reuben. 2004. The causes and consequences of residential segregation: an equilibrium analysis of neighborhood sorting. Technical report Working Paper.
- Beber, Bernd, Philip Roessler and Alexandra Scacco. 2014. "Intergroup violence and political attitudes: evidence from a dividing Sudan." *The Journal of Politics* 76(3):649–665.

- Beelmann, Andreas and Kim Sarah Heinemann. 2014. "Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs." *Journal of Applied Developmental Psychology* 35(1):10–24.
- Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the Royal statistical society: series B (Methodological)* 57(1):289–300.
- Bigler, Rebecca S and Lynn S Liben. 1993. "A cognitive-developmental approach to racial stereotyping and reconstructive memory in Euro-American children." *Child development* 64(5):1507–1518.
- Cameron, Jessica A, Jeannette M Alvarez, Diane N Ruble and Andrew J Fuligni. 2001. "Children's lay theories about ingroups and outgroups: Reconceptualizing research on prejudice." *Personality and Social Psychology Review* 5(2):118–128.
- Cameron, Lindsey and Adam Rutland. 2006. "Extended contact through story reading in school: Reducing children's prejudice toward the disabled." *Journal of Social Issues* 62(3):469–488.
- Campbell, Andrea Louise. 2003. *How policies make citizens: Senior political activism and the American welfare state*. Princeton University Press.
- Carrell, Scott E, Mark Hoekstra and James E West. 2015. The impact of intergroup contact on racial attitudes and revealed preferences. Technical report National Bureau of Economic Research.
- Carvacho, Héctor, Andreas Zick, Andrés Haye, Roberto González, Jorge Manzi, Caroline Kocik and Melanie Bertl. 2013. "On the relation between social class and prejudice: The roles of education, income, and ideological attitudes." *European Journal of Social Psychology* 43(4):272–285.
- Christ, Oliver, Miles Hewstone, Nicole Tausch, Ulrich Wagner, Alberto Voci, Joanne Hughes and Ed Cairns. 2010. "Direct contact as a moderator of extended contact effects: Cross-sectional and longitudinal impact on outgroup attitudes, behavioral intentions, and attitude certainty." *Personality and Social Psychology Bulletin* 36(12):1662–1674.
- Clayton, Katherine, Jeremy Ferwerda and Yusaku Horiuchi. 2018. "Exposure to Immigration and Admission Preferences: Evidence From France." *Available at SSRN 3268904* .
- Davies, Kristin, Linda R Tropp, Arthur Aron, Thomas F Pettigrew and Stephen C Wright. 2011. "Cross-group friendships and intergroup attitudes: A meta-analytic review." *Personality and Social Psychology Review* 15(4):332–351.
- Doerr, Celeste, E Ashby Plant, Jonathan W Kunstman and David Buck. 2011. "Interactions in Black and White: Racial differences and similarities in response to interracial interactions." *Group Processes & Intergroup Relations* 14(1):31–43.
- Dovidio, John F, Anja Eller and Miles Hewstone. 2011. "Improving intergroup relations through direct, extended and other forms of indirect contact." *Group processes & intergroup relations* 14(2):147–160.

- Doyle, Anna Beth and Frances E Aboud. 1995. "A longitudinal study of White children's racial prejudice as a social-cognitive development." *Merrill-Palmer Quarterly* (1982-) pp. 209–228.
- Enos, Ryan D. 2014. "Causal effect of intergroup contact on exclusionary attitudes." *Proceedings of the National Academy of Sciences* 111(10):3699–3704.
- Enos, Ryan D and Noam Gidron. 2016. "Intergroup behavioral strategies as contextually determined: Experimental evidence from Israel." *The Journal of Politics* 78(3):851–867.
- Epstein, Ralph and SS Komorita. 1966. "Childhood prejudice as a function of parental ethnocentrism, punitiveness, and outgroup characteristics." *Journal of Personality and Social Psychology* 3(3):259.
- Eric Oliver, J and Janelle Wong. 2003. "Intergroup prejudice in multiethnic settings." *American journal of political science* 47(4):567–582.
- Fearon, James D and David D Laitin. 2000. "Violence and the social construction of ethnic identity." *International organization* 54(4):845–877.
- Ferwerda, Jeremy, DJ Flynn and Yusaku Horiuchi. 2017. "Explaining opposition to refugee resettlement: The role of NIMBYism and perceived threats." *Science advances* 3(9):e1700812.
- Gelman, Andrew and John Carlin. 2014. "Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors." *Perspectives on Psychological Science* 9(6):641–651.
- Gil-White, Francisco Jose. 2001. "Are ethnic groups biological ?species? to the human brain? Essentialism in our cognition of some social categories." *Current anthropology* 42(4):515–553.
- Gómez, Angel, Linda R Tropp and Saulo Fernández. 2011. "When extended contact opens the door to future contact: Testing the effects of extended contact on attitudes and intergroup expectancies in majority and minority groups." *Group Processes & Intergroup Relations* 14(2):161–173.
- Graf, Sylvie, Stefania Paolini and Mark Rubin. 2014. "Negative intergroup contact is more influential, but positive intergroup contact is more common: Assessing contact prominence and contact prevalence in five Central European countries." *European Journal of Social Psychology* 44(6):536–547.
- Green, Eva GT, Oriane Sarrasin, Robert Baur and Nicole Fasel. 2016. "From stigmatized immigrants to radical right voting: A multilevel study on the role of threat and contact." *Political Psychology* 37(4):465–480.
- Habyarimana, James, Macartan Humphreys, Daniel N Posner and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101(4):709–725.
- Hassan, Adeel. 2019. "Hate-Crime Violence Hits 16-Year High, F.B.I. Reports." *The New York Times* .
- Hopkins, Daniel J. 2010. "Politicized places: Explaining where and when immigrants provoke local opposition." *American political science review* 104(1):40–60.

- Ioannou, Maria, Miles Hewstone and Ananthi Al Ramiah. 2017. "Inducing similarities and differences in imagined contact: A mutual intergroup differentiation approach." *Group Processes & Intergroup Relations* 20(4):427–446.
- Jones, Bradley. 2019. "Democrats far more likely than Republicans to see discrimination against blacks, not whites." *Pew Research Center* .
- Killen, Melanie and Adam Rutland. 2011. *Children and social exclusion: Morality, prejudice, and group identity*. John Wiley & Sons.
- Killen, Melanie, Adam Rutland and Martin D Ruck. 2011. "Promoting equity, tolerance, and justice in childhood." *Social Policy Report* 25(4):1–33.
- Kunovich, Robert M and Randy Hodson. 2002. "Ethnic diversity, segregation, and inequality: A structural model of ethnic prejudice in Bosnia and Croatia." *The Sociological Quarterly* 43(2):185–212.
- Lam, Chun Bun, Anna R Solmeyer and Susan M McHale. 2012. "Sibling relationships and empathy across the transition to adolescence." *Journal of youth and adolescence* 41(12):1657–1670.
- Lerman, Amy E and Katherine T McCabe. 2017. "Personal experience and public opinion: a theory and test of conditional policy feedback." *The Journal of Politics* 79(2):624–641.
- Lerman, Amy E and Vesla M Weaver. 2014. *Arresting citizenship: The democratic consequences of American crime control*. University of Chicago Press.
- Levy, Sheri R and Melanie Killen. 2008. *Intergroup attitudes and relations in childhood through adulthood*. Oxford University Press.
- Lowe, Matt. 2020. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *MIT Working Paper* .
- Mehrabian, Albert. 2017. *Nonverbal communication*. Routledge.
- Mettler, Suzanne and Andrew Milstein. 2007. "American political development from citizens' perspective: Tracking federal government's presence in individual lives over time." *Studies in American Political Development* 21(1):110–130.
- Milner, David. 1983. *Children & Race*. ERIC.
- Nasie, Meytal and Gil Diesendruck. 2019. "What children want to know about in-and out-groups, and how knowledge affects their intergroup attitudes." *Social Development* .
- Nipedal, Christian, Drew Nesdale and Melanie Killen. 2010. "Social group norms, school norms, and children's aggressive intentions." *Aggressive Behavior: Official Journal of the International Society for Research on Aggression* 36(3):195–204.

- Orfield, Gary, Erica Frankenberg, Jongyeon Ee and John Kuscera. 2014. "Great progress, a long retreat and an uncertain future." *UCLA: The Civil Rights Project*. Retrieved October 5:2014.
- Paluck, Elizabeth Levy. 2009. "Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda." *Journal of personality and social psychology* 96(3):574.
- Paluck, Elizabeth Levy. 2011. "Peer pressure against prejudice: A high school field experiment examining social network change." *Journal of Experimental Social Psychology* 47(2):350–358.
- Paluck, Elizabeth Levy, Seth Green and Donald P Green. 2017. "The Contact Hypothesis Revisited." *SSRN* .
- Patterson, Brenda. 2019. "Video Calls Fast Becoming as Popular as Voice Calls, Reaching Almost Universal Adoption for Social Use, According to Vonage Study." *The Business Wire* .
- Pettigrew, Thomas F. 1997. "Generalized intergroup contact effects on prejudice." *Personality and social psychology bulletin* 23(2):173–185.
- Pettigrew, Thomas F and Linda R Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of personality and social psychology* 90(5):751.
- Pettigrew, Thomas F and Linda R Tropp. 2013. *When groups meet: The dynamics of intergroup contact*. Psychology Press.
- Pettigrew, Thomas F, Oliver Christ, Ulrich Wagner and Jost Stellmacher. 2007. "Direct and indirect intergroup contact effects on prejudice: A normative interpretation." *International Journal of intercultural relations* 31(4):411–425.
- Putnam, Robert. 1993. "The prosperous community: social capital and public life." *The american prospect* 13(Spring), Vol. 4. Available online: <http://www.prospect.org/print/vol/13> (accessed 7 April 2003).
- Putnam, Robert D. 2007. "E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture." *Scandinavian political studies* 30(2):137–174.
- Raabe, Tobias and Andreas Beelmann. 2011. "Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences." *Child development* 82(6):1715–1737.
- Rao, Gautam et al. 2013. "Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools." *Job Market Paper* .
- Rizzolatti, Giacomo and Laila Craighero. 2004. "The mirror-neuron system." *Annu. Rev. Neurosci.* 27:169–192.
- Rothstein, Bo. 2005. *Social traps and the problem of trust*. Cambridge University Press.
- Rowley, Stephanie J, Margaret R Burchinal, Joanne E Roberts and Susan A Zeisel. 2008. "Racial identity, social context, and race-related social cognition in African Americans during middle childhood." *Developmental Psychology* 44(6):1537.

- Saguy, Tamar, John F Dovidio and Felicia Pratto. 2008. "Beyond contact: Intergroup contact in the context of power relations." *Personality and Social Psychology Bulletin* 34(3):432–445.
- Saguy, Tamar, Linda R Tropp and Diala Hawi. 2012. The role of group power in intergroup contact. In *Advances in intergroup contact*. Psychology Press pp. 127–146.
- Scacco, Alexandra and Shana S Warren. 2018. "Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria." *American Political Science Review* pp. 1–24.
- Schaffner, B. 2018. "Follow the racist? The consequences of Trump's expressions of prejudice for mass rhetoric." *Unpublished paper* .
- Schroeder, Juliana, Jane Risen, Francesca Gino and Michael Norton. 2014. "Handshaking promotes cooperative dealmaking."
- Simonivitz, Gabor, Gabor Kezdi and Peter Kardos. 2017. "Seeing the World Through the Other's Eye: An Online Intervention Reducing Ethnic Prejudice." *American Political Science Review* pp. 1–8.
- Skinner, Allison L, Andrew N Meltzoff and Kristina R Olson. 2017. "Catching? social bias: Exposure to biased nonverbal signals creates social biases in preschool children." *Psychological Science* 28(2):216–224.
- Sniderman, Paul. 2015. "Myrdal's Insight." *Mimeo* .
- Sniderman, Paul M and Philip E Tetlock. 1986. "Symbolic racism: Problems of motive attribution in political analysis." *Journal of social issues* 42(2):129–150.
- Sønderskov, Kim Mannemar and Jens Peter Frølund Thomsen. 2015. "Contextualizing intergroup contact: do political party cues enhance contact effects?" *Social Psychology Quarterly* 78(1):49–76.
- Soss, Joe. 2002. *Unwanted claims: The politics of participation in the US welfare system*. University of Michigan Press.
- Stathi, Sofia, Lindsey Cameron, Bonny Hartley and Shona Bradford. 2014. "Imagined contact as a prejudice-reduction intervention in schools: The underlying role of similarity and attitudes." *Journal of Applied Social Psychology* 44(8):536–546.
- Statistics Canada. 2017. "Canada (table). Census profile. 2016 Census. Statistics Canada catalogue no. 98-316-X2016001."
- Strauss, Valerie. 2020. "Millions of students could be home for the rest of academic year because of coronavirus, officials warn." *The Washington Post* .
- Todd, Andrew R and Adam D Galinsky. 2014. "Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications." *Social and Personality Psychology Compass* 8(7):374–387.
- Tropp, Linda R and Mary A Prenovost. 2008. "The role of intergroup contact in predicting children's interethnic attitudes: Evidence from meta-analytic and field studies."

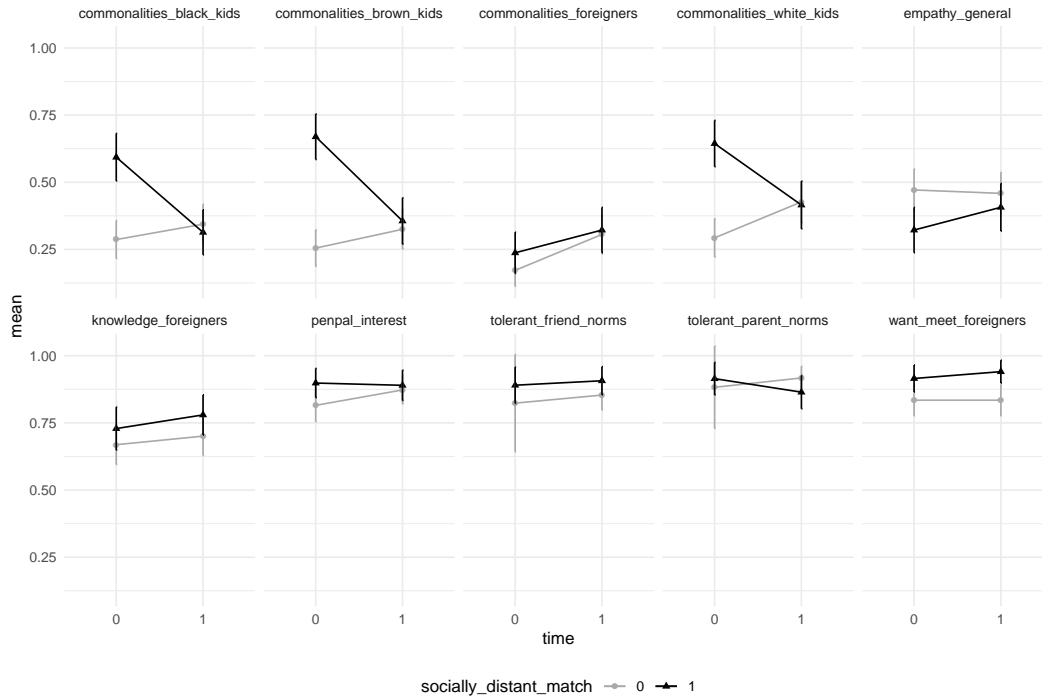
- Tropp, Linda R, Thomas C O'Brien and Katya Migacheva. 2014. "How peer norms of inclusion and exclusion predict children's interest in cross-ethnic friendships." *Journal of Social Issues* 70(1):151–166.
- Tropp, Linda R, Thomas C O'Brien, Roberto González Gutierrez, Daniel Valdenegro, Katya Migacheva, Pablo de Tezanos-Pinto, Christian Berger and Oscar Cayul. 2016. "How school norms, peer norms, and discrimination predict interethnic experiences among ethnic minority and majority youth." *Child Development* 87(5):1436–1451.
- Turner, Rhiannon N, Miles Hewstone, Alberto Voci and Christiana Vonofakou. 2008. "A test of the extended intergroup contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self." *Journal of personality and social psychology* 95(4):843.
- Vezzali, Loris, Dino Giovannini and Dora Capozza. 2012. "Social antecedents of children's implicit prejudice: Direct contact, extended contact, explicit and implicit teachers' prejudice." *European Journal of Developmental Psychology* 9(5):569–581.
- Wagner, Ulrich, Rolf Van Dick, Thomas F Pettigrew and Oliver Christ. 2003. "Ethnic prejudice in East and West Germany: The explanatory power of intergroup contact." *Group Processes & Intergroup Relations* 6(1):22–36.
- West, Martin R. 2016. "Should non-cognitive skills be included in school accountability systems? Preliminary evidence from California's CORE districts." *Evidence Speaks Reports* 1(13):1–7.
- West, Tessa V, Joe C Magee, Sarah H Gordon and Lindy Gullett. 2014. "A little similarity goes a long way: The effects of peripheral but self-revealing similarities on improving and sustaining interracial relationships." *Journal of personality and social psychology* 107(1):81.
- Wright, Stephen C, Arthur Aron, Tracy McLaughlin-Volpe and Stacy A Ropp. 1997. "The extended contact effect: Knowledge of cross-group friendships and prejudice." *Journal of Personality and Social psychology* 73(1):73.



# Appendix

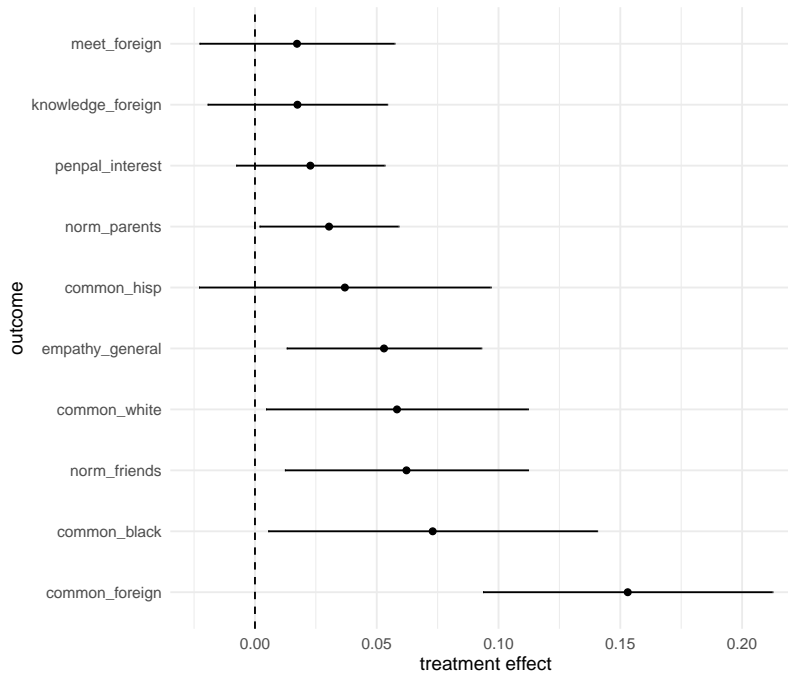
## Additional Analyses

**Figure 9: Raw Means: Socially Distant Matches vs. Control Group ( $n = 508$ )**



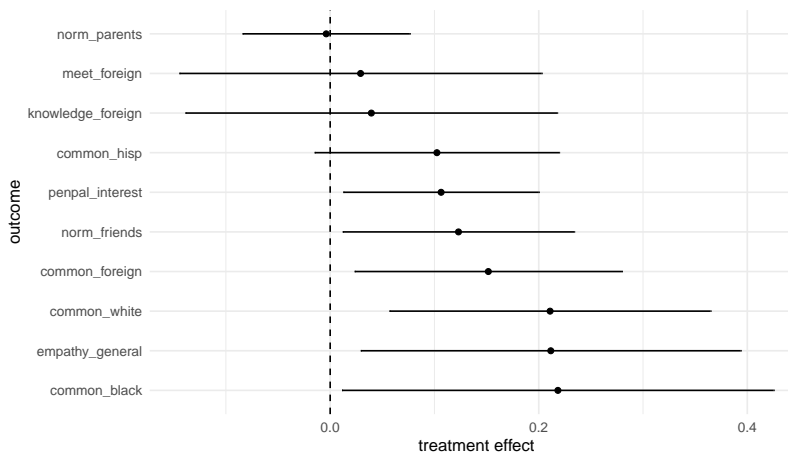
Outcomes are binary and coded in a tolerant direction. Error bars represent 90% confidence intervals.

**Figure 10: ATE: Binary Outcomes ( $n = 508$ )**



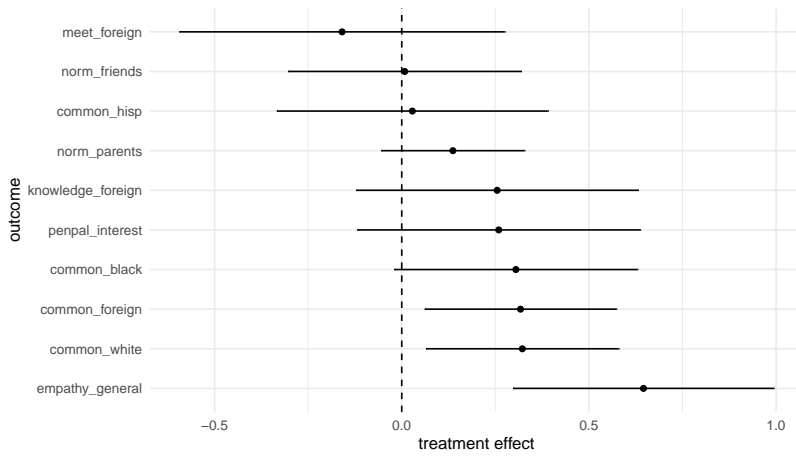
Outcomes are binary and coded in a tolerant direction. Estimates are based on an OLS model controlling for treatment block, student ethnicity, language spoken, having siblings, and country. Standard errors are clustered at the classroom level. Error bars represent 90% confidence intervals.

**Figure 11: Average Treatment Effects: White Children ( $n = 412$ )**



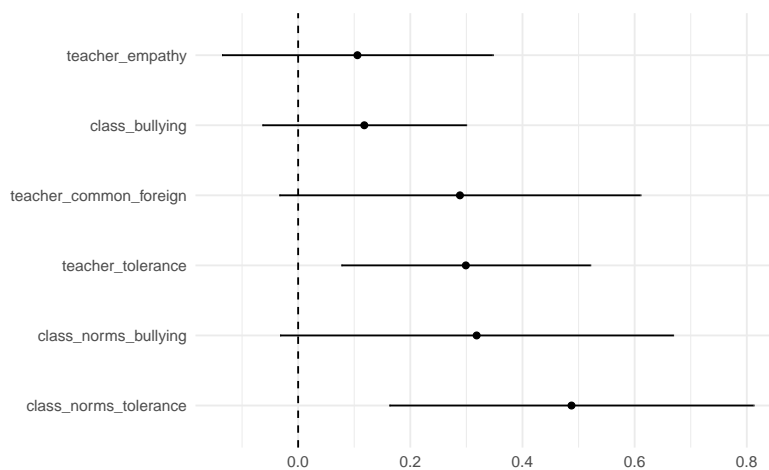
Sample subset to children who self-identify as White ( $n = 403$ ). Treatment coefficients based on OLS regressions where outcomes are continuous, and differenced-out ( $t_1 - t_0$ ). Models include the following covariates: randomization block, outcome measured at baseline, language spoken, country, and if one has a sibling.

**Figure 12: Average Treatment Effects: Non-White Children ( $n = 91$ )**



Sample subset to children who self-identify as non-White ( $n = 91$ ). Treatment coefficients based on OLS regressions where outcomes are continuous, and differenced-out ( $t_1 - t_0$ ). Models include the following covariates: randomization block, outcome measured at baseline, language spoken, country, and if one has a sibling.

**Figure 13: Over-Time Changes: Treated Teachers ( $n = 87$ )**



Outcomes reflect teacher-reported changes to rates of classroom bullying, norms around intervening when a student is being bullied, and norms around tolerating children of different socio-economic and ethnic backgrounds

**Table 1:** ATE: Difference-in-Differences

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.021 (0.085)	-0.029 (0.095)	0.128* (0.067)	0.101* (0.060)	0.015 (0.050)	0.090 (0.062)	0.143 (0.088)	0.198* (0.102)	0.113 (0.081)	0.259*** (0.073)
Constant	0.304 (0.341)	0.718 (0.586)	-0.550* (0.306)	3.499*** (0.305)	3.538*** (0.356)	0.660 (0.586)	0.704 (0.609)	-1.371*** (0.426)	-1.543*** (0.505)	-2.860*** (0.419)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.073	0.062	0.162	0.214	0.194	0.110	0.139	0.409	0.401	0.433
Adjusted R <sup>2</sup>	-0.0002	-0.012	0.096	0.152	0.130	0.040	0.071	0.362	0.354	0.389

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

These estimates are based on OLS regressions where outcomes are continuous, and differenced-out ( $t_1 - t_0$ ). Models include the following covariates: randomization block, outcome measured at baseline, language spoken, country, if one has a sibling, and ethnicity (proxied by self-reported skin tone). The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 2:** ATE: Logistic Regressions

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Blk.	Common Brn.	Common Wht.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.076 (0.154)	0.317 (0.272)	1.042*** (0.246)	0.581** (0.277)	0.415* (0.227)	0.326* (0.190)	0.276** (0.129)	0.426* (0.243)	0.191 (0.198)	0.294* (0.167)
Constant	5.340*** (1.604)	12.031*** (1.880)	-7.878*** (1.484)	12.509*** (1.664)	23.846*** (1.175)	12.115*** (2.085)	-6.771*** (1.857)	-6.972*** (1.537)	-7.339*** (1.421)	-11.321*** (1.757)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.321	0.275	0.261	0.255	0.294	0.225	0.281	0.252	0.218	0.237
χ <sup>2</sup>	128.211***	80.060***	98.249***	80.758***	80.833***	63.244***	119.627***	98.766***	85.986***	97.944***

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

These estimates are based on logistic regressions with the following covariates: randomization block, outcome measured at baseline, language spoken, country, if one has a sibling, and ethnicity (proxied by self-reported skin tone). The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 3: ATE: Binary Outcomes using OLS**

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.017 (0.022)	0.017 (0.024)	0.153*** (0.036)	0.062** (0.030)	0.030* (0.017)	0.023 (0.019)	0.053** (0.024)	0.073* (0.041)	0.037 (0.036)	0.058* (0.033)
Constant	0.414** (0.172)	1.078*** (0.123)	-0.014 (0.226)	1.173*** (0.153)	1.298*** (0.121)	1.001*** (0.135)	0.191 (0.225)	-0.131 (0.250)	-0.024 (0.209)	-0.649*** (0.177)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.243	0.164	0.207	0.145	0.160	0.121	0.215	0.188	0.165	0.179
Adjusted R <sup>2</sup>	0.182	0.096	0.142	0.078	0.094	0.050	0.152	0.122	0.097	0.112

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

These estimates are based on OLS regressions where outcomes are continuous. Models include the following covariates: randomization block, outcome measured at baseline, language spoken, country, if one has a sibling, and ethnicity (proxied by self-reported skin tone). The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 4: Heterogenous Effects by Social Distance**

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Blk.	Common Brn.	Common Wht.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.006 (0.047)	0.027 (0.055)	0.206*** (0.058)	0.012 (0.046)	0.076** (0.033)	0.015 (0.031)	0.182*** (0.023)	0.201*** (0.060)	0.112*** (0.042)	0.149*** (0.056)
Distant Match	0.077 (0.087)	0.144 (0.069)	0.012 (0.098)	0.078 (0.084)	0.009 (0.060)	0.313*** (0.043)	0.042 (0.042)	0.075 (0.071)	-0.058 (0.061)	(0.057)
Treated : Distant Match	0.018 (0.049)	-0.016 (0.062)	-0.081 (0.070)	0.078 (0.057)	-0.069* (0.036)	0.012 (0.035)	-0.197*** (0.037)	-0.197*** (0.071)	-0.116** (0.058)	-0.138* (0.071)
Constant	0.443*** (0.167)	0.999*** (0.140)	-0.170 (0.188)	1.173*** (0.122)	1.209*** (0.088)	0.994*** (0.132)	-0.148 (0.188)	-0.199 (0.220)	0.067 (0.230)	-0.612*** (0.155)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.243	0.164	0.208	0.147	0.163	0.121	0.223	0.197	0.168	0.183
Adjusted R <sup>2</sup>	0.180	0.094	0.143	0.078	0.095	0.048	0.158	0.130	0.099	0.115

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Models are the same as those used for the main OLS specification (10) with the addition of an interaction term to reflect social distance, operationalized using a binary indicator for whether the exchange involved a partner classroom that either shared over 25% of the same ethnicity, or was located in a foreign country. Models 1 — 3 capture attitudes toward foreigners. The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 5: Heterogenous Effects by Minority Status**

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.055 (0.048)	0.053 (0.045)	0.200*** (0.070)	0.051 (0.059)	0.072** (0.032)	0.025 (0.037)	0.111** (0.049)	0.134** (0.059)	0.085 (0.056)	0.059 (0.077)
Minority	0.014 (0.086)	0.059 (0.054)	-0.015 (0.044)	-0.047 (0.056)	0.066 (0.062)	0.046 (0.056)	0.004 (0.084)	-0.017 (0.075)	0.079 (0.100)	0.004 (0.068)
Treated: Minority	-0.065 (0.071)	-0.063 (0.060)	-0.079 (0.079)	0.021 (0.080)	-0.075 (0.050)	-0.005 (0.056)	-0.098 (0.068)	-0.104 (0.079)	-0.086 (0.075)	-0.002 (0.099)
Constant	0.478*** (0.171)	1.134*** (0.109)	-0.064 (0.231)	1.064*** (0.156)	1.229*** (0.116)	1.099*** (0.139)	0.269 (0.207)	0.072 (0.230)	0.034 (0.178)	-0.454*** (0.175)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.239	0.160	0.200	0.144	0.153	0.111	0.215	0.191	0.164	0.173
Adjusted R <sup>2</sup>	0.177	0.092	0.135	0.077	0.086	0.039	0.151	0.125	0.096	0.106

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

OLS models are the same as those used for the main OLS specification (10) with the addition of an interaction term to reflect the respondent's minority status, operationalized using a binary indicator for whether the respondent does not belong to the majority ethnicity in the classroom. Models 1 — 3 capture attitudes toward foreigners. The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.



**Table 6:** Heterogenous Effects for White-Majority Classrooms

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.006 (0.024)	0.035 (0.028)	0.143*** (0.037)	0.079** (0.034)	0.015 (0.019)	0.044** (0.020)	0.019 (0.030)	0.065 (0.053)	0.039 (0.045)	0.061** (0.030)
White	-0.033 (0.100)	0.023 (0.107)	0.054 (0.073)	-0.012 (0.092)	-0.096 (0.080)	0.063 (0.039)	0.103 (0.076)	-0.004 (0.112)	-0.061 (0.071)	0.318* (0.183)
Treated: White	0.036 (0.053)	-0.057 (0.053)	0.033 (0.095)	-0.053 (0.067)	0.050 (0.035)	-0.067** (0.030)	0.108*** (0.040)	0.026 (0.079)	-0.007 (0.073)	-0.008 (0.099)
Constant	0.425** (0.174)	1.062*** (0.129)	-0.003 (0.230)	1.158*** (0.155)	1.312*** (0.125)	0.985*** (0.135)	0.222 (0.227)	-0.123 (0.251)	-0.026 (0.207)	-0.651*** (0.178)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.244	0.166	0.210	0.146	0.164	0.123	0.218	0.193	0.167	0.179
Adjusted R <sup>2</sup>	0.180	0.095	0.143	0.075	0.094	0.048	0.151	0.124	0.096	0.108

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Models are the same as those used for the main OLS specification (10) with the addition of an interaction term to reflect if the respondent belongs to a majority (over 50%) White classroom. Models 1 — 3 capture attitudes toward foreigners. The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 7: Heterogenous Effects By Ethnicity**

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.100 (0.080)	0.066 (0.068)	0.211** (0.089)	-0.002 (0.079)	0.057 (0.050)	0.054 (0.067)	0.070 (0.103)	0.118 (0.116)	0.092 (0.104)	0.206** (0.103)
White	-0.013 (0.085)	0.056 (0.067)	0.051 (0.068)	-0.063 (0.049)	0.046 (0.057)	0.124** (0.063)	-0.003 (0.092)	-0.094 (0.091)	-0.074 (0.091)	0.236*** (0.079)
Treated: White	-0.096 (0.098)	-0.062 (0.080)	-0.076 (0.095)	0.075 (0.095)	-0.037 (0.063)	-0.044 (0.079)	-0.025 (0.110)	-0.057 (0.121)	-0.062 (0.108)	-0.174 (0.120)
Constant	0.594*** (0.142)	1.027*** (0.109)	-0.304* (0.173)	1.003*** (0.147)	1.068*** (0.104)	0.917*** (0.128)	0.112 (0.126)	-0.043 (0.156)	0.185 (0.165)	-0.295** (0.146)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.241	0.162	0.201	0.145	0.153	0.111	0.213	0.188	0.162	0.177
Adjusted R <sup>2</sup>	0.180	0.094	0.136	0.078	0.086	0.039	0.149	0.122	0.094	0.111

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

OLS models are the same as those used for the main OLS specification (10) with the addition of an interaction term to reflect the respondent's minority status, operationalized using a binary indicator for whether the respondent does not belong to the majority ethnicity in the classroom. Models 1 — 3 capture attitudes toward foreigners. The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

**Table 8:** Heterogenous Effects by Multi-Class Teacher Status

	<i>Outcome</i>									
	Info	Meet	Foreign	Norm Friends	Norm Parents	Penpal	Empathy	Common Black	Common Brown	Common White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.017 (0.024)	0.011 (0.027)	0.174*** (0.038)	0.047 (0.030)	0.047*** (0.018)	0.024 (0.020)	0.063** (0.027)	0.099** (0.040)	0.038 (0.038)	0.066* (0.035)
Multi-Class	-0.130 (0.087)	0.016 (0.082)	0.861*** (0.108)	0.058 (0.068)	0.119** (0.047)	0.135*** (0.047)	0.417*** (0.063)	0.554*** (0.171)	0.459*** (0.125)	-0.172 (0.181)
Treated: Multi-Class	0.002 (0.064)	0.044 (0.045)	-0.142* (0.085)	0.107 (0.073)	-0.113*** (0.034)	-0.008 (0.024)	-0.070 (0.060)	-0.182* (0.109)	-0.009 (0.084)	-0.051 (0.099)
Constant	0.473*** (0.173)	1.152*** (0.116)	-0.123 (0.222)	1.084*** (0.161)	1.204*** (0.106)	1.110*** (0.137)	0.237 (0.209)	0.003 (0.222)	0.039 (0.177)	-0.467*** (0.169)
Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	508	508	508	508	508	508	508	508	508	508
R <sup>2</sup>	0.243	0.164	0.210	0.148	0.164	0.121	0.216	0.193	0.165	0.179
Adjusted R <sup>2</sup>	0.180	0.094	0.144	0.079	0.097	0.048	0.151	0.126	0.095	0.111

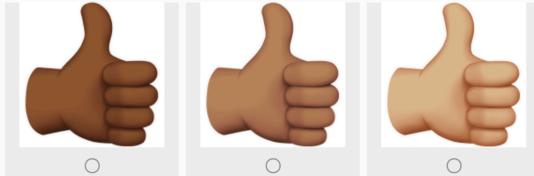
*Note:* \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

OLS models are the same as those used for the main OLS specification (10) with the addition of an interaction term to reflect the respondent's minority status, operationalized using a binary indicator for whether the respondent does not belong to the majority ethnicity in the classroom. Models 1 — 3 capture attitudes toward foreigners. The outcomes from left to right are: knowledge of life in other countries, interest in meeting a foreign child, perceived commonalities with foreigners, whether one's friends would accept a cross-group friendship, whether one's parents would accept a cross-group friendship, interest in having a foreign penpal, general empathy and perspective-taking, and perceived commonalities with Black children, Brown children, and White children.

## Survey Items

**Figure 14: Skin Tone**


Which of these pictures is the closest to your skin color?



Three thumbs-up emojis are displayed side-by-side, each with a different skin tone. From left to right, the skin tones are dark brown, medium brown, and light brown. Below each emoji is a small, empty radio button for selection.

I use this item as a proxy for ethnicity. This item is asked last in the endline survey to avoid priming effects.


**Figure 15: Commonalities with White Children**



How much do you think you have in common with the kids in this picture?

A photograph of two white children, a boy and a girl, smiling and standing outdoors. The boy is on the left, wearing a red and green plaid shirt. The girl is on the right, wearing a white cable-knit sweater. They are standing in front of a wooden fence and green foliage.

**Figure 16: Commonalities with Black Children**



How much do you think you have in common with the kids in this picture?

A photograph of two Black children, a boy and a girl, smiling and standing outdoors. The boy is on the left, wearing an orange shirt. The girl is on the right, wearing a denim jacket over a purple shirt. They are standing in front of green foliage.

**Figure 17:** Commonalities with Brown Children

