Broaching prejudicial taboos to improve intergroup relations:

A field experiment in Israel [†]

Chagai M. Weiss[‡] Shira Ran [§] Eran Halperin[¶]

March 21, 2022

Abstract

Prejudice reduction interventions often focus on intergroup similarities, positive shared experiences, and cooperation, while avoiding explicit engagement with prejudicial taboos relating to negative stereotypes, dehumanizing beliefs, and contentious political disagreements. Since prejudicial taboos are an unspoken foundation of intergroup tensions, we argue that interventions which overlook them, and fail to address the "elephant in the room," might fall short of challenging the foundations of intergroup prejudice. Therefore, we suggest that constructively broaching prejudicial taboos could be an effective way to improve intergroup relations. Building on this argument we develop a scalable educational program in which elementary school students watch and discuss episodes of an Israeli TV series "You Can't Ask That," which depicts children from minority groups responding to taboo questions. To test the effects of our intervention, we implemented a field-experiment in Israel which coincided with a cycle intense inter-communal violence. Our findings suggest that broaching prejudicial taboos affected attitudes and behaviors. Compared with students in a control group, treated students' positive affect towards outgroups increased by over 0.38 SD, their perceptions of intergroup similarity increased by over 0.36 SD, and their willingness to sign up for an intergroup contact initiative increased by 0.17 SD, 7-14 days post-treatment. We demonstrate that these effects are driven by attitudinal change amongst prejudicial students, and emphasize how constructively addressing prejudicial taboos head on can be an effective strategy for prejudice reduction, especially amongst prejudicial individuals, and even in times of intense intergroup violence.

[†]This study was pre-registered on OSF (https://osf.io/kdt8y), and was approved by the IRB office at The Hebrew University of Jerusalem, as well as by Israel's Ministry of Education. Achord center, and specifically Ronit Hanzis, Ido Oren, and Shir Tankel provided excellent support in designing and implementing the intervention. We thank participants at the MEI workshop at the Harvard Kennedy School for helpful comments and suggestions.

[‡]PhD candidate in Political Science, University of Wisconsin – Madison; Middle East Initiative Predoctoral Fellow, Harvard Kennedy School. Email: cmweiss@wisc.edu

[§]PhD candidate in Psychology, Hebrew University of Jerusalem.
Email: shira.ran@mail.huji.ac.il

[¶]Professor of Psychology, Hebrew University of Jerusalem.
Email: eran.halperin@mail.huji.ac.il

Prejudicial taboos which include negative stereotypes, dehumanizing beliefs, and contentious political disagreements, are a central component of intergroup tensions. For example, many Israelis believe that "a culture of violence" explains the prevalence of crime in Arab communities (Smooha 2015), and many white Americans believe that "a disposition towards laziness" explains black Americans economic disadvantage (Kinder, Sanders and Sanders 1996; Kam and Burge 2018). Despite the prevalence of prejudicial taboos, they often remain unspoken in intergroup interactions, but influential in shaping intergroup perceptions and policy preferences (Weber et al. 2014). Moreover, existing approaches for prejudice reduction rarely broach these prejudicial taboos directly, in an attempt to improve intergroup relations (Saguy et al. 2009; Saguy 2018; Kteily and McClanahan 2020).

Avoiding sensitive topics, existing approaches for prejudice reduction often focus on harmonious interventions relating to cooperation (Scacco and Warren 2018; Mousa 2020), intergroup similarities (Williamson et al. 2021), norms of inclusion (Murrar, Campbell and Brauer 2020), and empathy building (Kalla and Broockman 2021). While some of these approaches have proven remarkably effective in shaping attitudes and behaviors (Broockman and Kalla 2016; Murrar, Campbell and Brauer 2020; Kalla and Broockman 2021), recent meta-analyses emphasize the modest effects of most interventions, and point to theoretical and empirical limitations of the prejudice reduction literature, calling researchers to develop and test novel theoretically-informed interventions (Paluck, Green and Green 2019; Paluck et al. 2021). Motivated by this call, and by the overwhelming focus of existing research on interventions which avoid direct engagement with sensitive topics, we examine an alternative approach, and test how constructively broaching prejudicial taboos affects intergroup prejudice.

Determining the effectiveness of prejudice reduction interventions which tackle prejudicial taboos directly is an important endeavor for social scientists, with direct practical implications. Avoiding direct engagement with prejudicial taboos might make inter-group exchanges less threatening and more inviting for majority group members (Ron et al. 2017), especially if social norms and personal motivations for political correctness encourage people to shy away from taboos

(Langer et al. 1976; Weber et al. 2014). However, such interventions which overlook the deep roots and uncomfortable foundations of intergroup prejudice, might fail to address "the elephant in the room," and fall short of challenging the foundations of intergroup tensions.

In what follows we lay out a novel approach for prejudice reduction, which broaches prejudicial taboos directly. Doing so, we attempt to challenge prevailing notions in the literature on intergroup relations, by which prejudice reduction requires harmonious and cooperative interactions between majority and minority group members (Allport 1954). To test the effectiveness of interventions that broach taboos directly, we designed a month-long educational intervention based on the Israeli TV-series "You Can't Ask That."[1] In each episode of the TV-series, members of a different minority group are filmed answering taboo questions sourced from home audiences (see Figure 1). Our educational intervention, focused on three episodes depicting Arab, visually impaired, and immigrant children,[2] and included four weekly sessions in which Jewish elementary school children watched the TV-series and engaged in guided follow-up classroom discussions. The core objective of the intervention was to broach prejudicial taboos, and discuss them in a constructive fashion, in order to reduce prejudice towards multiple minorities, and foster more favorable intergroup attitudes and behaviors.

Broaching Taboos to Reduce Prejudice

Though the specific content of prejudicial taboos is context and time dependent, such taboos often include negative stereotypes, dehumanizing beliefs, and contentious political disagreements. The starting point of our intervention is that prejudicial taboos, which are rarely discussed in cross-grop interactions, are a foundational component of intergroup tensions. In line with this insight, existing studies suggest that humanizing outgroups (Haslam and Loughnan 2014), emphasizing similarities, and exposing people to counter stereotypical personalities or information can reduce

[1]The TV series is a Hebrew adaptation of an Australian show, which has been translated and aired in multiple countries including the U.S., Canada, and Belgium. See Israeli TV series website here: https://testkankids.kan.org.il/program/?catid=1527.

[2]Immigrants were children of Filipino foreign workers, many of whom are undocumented immigrants in Israel.

Figure 1: "You Can't Ask That" excerpt. This Figure depicts a moment from an episode focusing on Arab children, in which an Arab girl responds to a question asking "Why is Arabic a scary language?"

intergroup prejudice (Ramasubramanian 2011, 2015; Choi, Poertner and Sambanis 2021; MARBLE et al. 2021). However, doing so does not directly confront prevalent taboos. Instead, such approaches provide competing information to correct or update individual perceptions, and in doing so they might fall short of reducing prejudice, because of individual tendencies towards sub-typing (Richards and Hewstone 2001).

Existing approaches which provide competing information instead of broaching prejudicial taboos might be motivated by the fact that tackling sensitive topics head on is often uncomfortable (Ron et al. 2017). More so, engaging with taboos might possibly reinforce negative beliefs and perpetuate group differences. However, avoding direct engagement with taboos risks the possibility that they remain unchallenged. Thus one may wonder if and how explicitly and constructively broaching prevalent taboos might affect majority group members' prejudice towards minorities.

We argue that broaching prejudicial taboos, and exposing majority group members to minorities who constructively dismantle taboos, can improve intergroup attitudes and behaviors. Doing so addresses the uncomfortable albeit foundational elements of intergroup prejudice which often remain unaddressed in conventional interventions. More so, if done in a constructive way, the process of dismantling taboos can further activate three common mechanisms which are conducive for prejudice reduction.

First, since different elements of prejudicial taboos are often associated with perceptions of

group homogeneity (Fiske 1998), broaching taboos can increase people's perceptions of group variability (Brauer and Er-Rafiy 2011; Er-rafiy and Brauer 2013). For example, constructive discussions regarding contentious events of intergroup violence can emphasize that different outgroups hold varying political preferences, and that not all outgroups seek to hurt one's ingroup. In turn heightened perceptions of group variability, as a result of engagement with taboos, can serve to reduce prejudice (Hsieh, Faulkner and Wickes 2021). Second, when majority group members witness minorities directly addressing prejudicial taboos, they receive novel information about the outgroup. Such information can increase perceptions of intergroup similarity. For example, a critical discussion about citizenship status and national identification can serve to emphasize to natives that many immigrants identify with the nation and appreciate local culture. In turn realizations about intergorup similarity can improve majority group members attitudes towards minorities (Falomir-Pichastor, Martínez and Paterna 2010; Brandt 2017; Liberman, Woodward and Kinzler 2017). Finally, broaching taboos can provide majority group members with insights about the type of challenges that minorities face in their daily life. For example, explicitly discussing the content and prevalence of dehumanizing beliefs held by majority group members could emphasize how often minorities suffer from hostile interactions, and how hurtful such interactions may be. These insights about minority group members' challenges, provide a unique opportunity for perspective getting. In turn, experiences of perspective getting can serve to reduce prejudice (Bruneau and Saxe 2012; Kalla and Broockman 2021).

Taken together, our theoretical framework suggests that broaching prejudicial taboos, and addressing "the elephant in the room" in a constructive fashion, can be an effective way to reduce prejudice. In light of this insight, we designed our intervention to focus on three episodes of the TV series "You Can't Ask That." The main objective of the show is to provide home audiences with an opportunity to ask minorities forthright questions regarding prejudicial taboos, and to generate a constructive discussion (in the studio) about sensitive topics regarding intergroup relations. Thus, the show provides a novel platform to address some of the most sensitive and contentious topics relating to intergroup relations.

Our primary goal was to expose majority group students to episodes in which minority children respond to taboo questions, which many children are curious about, yet too shy to ask. The episodes we selected, focus on three outgroups: Arab, visually impaired, and immigrant children. Within these episodes, Arabs children were asked questions such as *Why is Arabic a scary language? Do you think we are enemies?* immigrant children were asked other questions, including *Do you eat weird foods? Do you think you are Israeli?* and visually impaired children were asked questions about their disability such as *What do you see? Do you get bullied regularly?* In response to these questions, children further discussed a host of complex and sensitive topics relating to group identity, political disagreement, experiences of racism and abselism, cultural differences, and physical disability.

Our intervention included four meetings, which were designed to constructively discuss prejudicial taboos. In the first three meetings, students watched a group-specific episode (15 minutes), and engaged in follow-up classroom discussions (30 minutes). In the final meeting, students watched a recap from all three episodes (15 minutes), and engaged in an overview discussion (30 minutes). We designed classroom discussions to focus on the TV series main theme—prejudicial taboos—and to connect this theme with the core theoretical mechanisms noted above, relating to: information about outgroup variability, intergroup similarities, and perspective getting. In doing so, we ensured to broach prejudicial taboos in a constructive fashion. An Elaborate description of the TV series is provided in Appendix S1, and an overview of our educational program is described in Appendix S2.

Research Design

To test our intervention we implemented a field experiment with 12 Israeli classes in grades 4-6. After receiving IRB approval from the Hebrew University of Jerusalem, permission from Israel's ministry of education, and informed consent from students' parents, our field experiment followed three steps depicted in Figure 2.[3] First, we surveyed 270 students from 12 classes in grades 4-6,

[3]In line with our IRB approval, and agreement with the ministry of education, all students in treated classes participated in our intervention, however only students for whom we received

to collect baseline demographics and pre-treatment measures of our main outcomes. After completing the baseline survey, we block randomized classes into treatment and control conditions by grade, resulting in 6 treated and 6 untreated classes. Students in treated classes, participated in our month-long educational curriculum, which was delivered by a professional educational practitioner.

As depicted in Figure 2, the start of our intervention coincided with a cycle of intense violence between Jews and Palestinians. Between May 10-21 2021, intense missile fires and intercommunal clashes disrupted life in many cities across Israel, including our intervention site. Violence was so intense, that some schools closed for several days, but during the study period, our partner school largely operated, and we concluded implementing our intervention amongst treated classes in the first week of June.

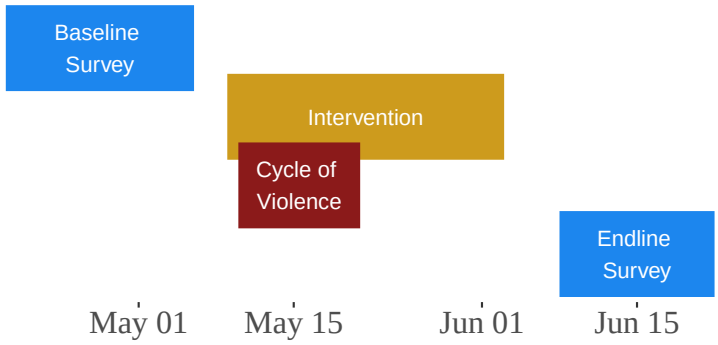

Figure 2: **Study Timeline**

A week post-treatment, we began collecting endline surveys from 253 students, and students were interviewed 1-2 weeks after the intervention was concluded. In our surveys, we measured attitudes and behaviors relating to outgroup affect, perceptions of similarity with outgroups, support for diversity, interest in intergroup contact, and curiosity about outgroups.We describe the survey wording we used to collect our main outcome measures in Appendix S3.

parental informed consent participated in our surveys. We obtained informed consent for over 70% of students.

Estimation Strategy

To identify the effects of our intervention, we estimate OLS regressions in which we interact mean centered covariates ($\boldsymbol{Z}_i$) with our main treatment indicator (X_i) (Lin 2013), and cluster errors by classroom (Abadie et al. 2017). Covariates include gender, assignment block, and pre-treatment outcome measures.[4] Our estimating equation is presented below:[5]

$$y_{ic} = \beta X_i + \phi \boldsymbol{Z}_i + \gamma(X_i * \boldsymbol{Z}_i) + \varepsilon_{ic} \tag{1}$$

Throughout all analyses, our interest is in identifying βX_i, representing the average treatment effect of the intervention on students' post treatment attitudes and behaviors.

Results

In Figure 3, we consider the effects of our intervention on students' positive affect and feeling of similarity towards outgroups. We report both group specific measures (e.g. students' positive affect towards Arabs), and aggregate scales measuring overall affect and similarity towards multiple outgroups combined. Each point estimate in Figure 3 represents the average treatment effect of our intervention on a standardized outcome of interest. We include measures related to groups mentioned in the intervention (e.g. Arab children), as well as Ultra-Orthodox people which were not mentioned in the intervention, in order to consider the extent to which our effects are group specific or more general.

As reported in the left panel of Figure 3, our treatment had a substantial effect on students' positive affect towards outgroups. Indeed, participating in the classroom intervention increased students' ranking of outgroups on a 0-100 feeling thermometer index by over 38% of a SD. In

[4]For our two behavioral measures which were not measured pre-treatment, we adjust our model with pre-treatment thermometer, diversity, similarity, and contact intention indices.

[5]The specification presented in Equation 1 is subtly different from our pre-registered specification, as further discussed in Appendix S9. Our results remain similar in our pre-registered models (where we control for pre-treatment covariates without interacting them with treatment), but we adapt an interaction specification following recent recommendations in the literature (Lin 2013). We report pre-registered models, as well as additional alternative specifications in Section S6 of the appendix.

practice, this effect resembles an 8 points shift on a feeling thermometer, approximately double the magnitude of the average effect of well powered intervention report in a recent meta analysis of prejudice reduction experiments (Paluck et al. 2021).

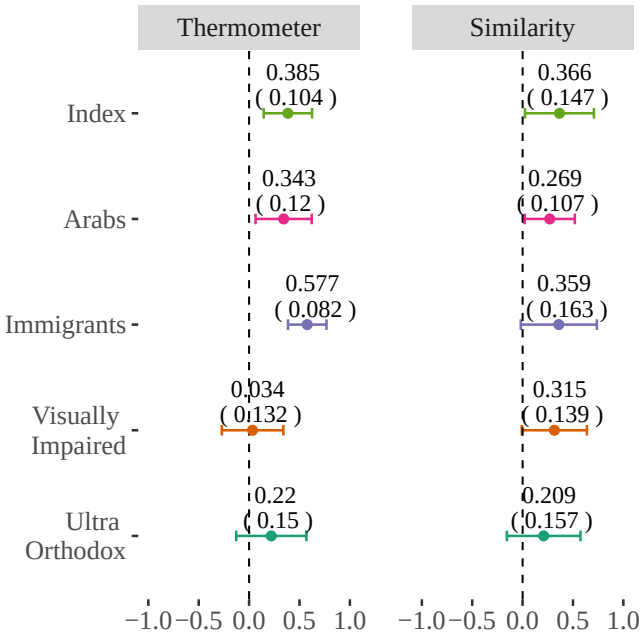

Figure 3: **Intervention effects on feeling thermometers and perceptions of outgroup similarity.** Each point estimates and corresponding 95% confidence interval is extracted from an OLS model specified in Equation 1.

When turning to examine group specific measures, some interesting patterns emerge. Focusing on groups discussed as part of the intervention, it appears that the treatment increased positive affect towards Arabs and immigrants, but had no effect on affect towards visually impaired children, likely because of ceiling effects in affect towards this group ($\mu = 75.9$, $\sigma^2 = 24.2$, in the pre-treatment period).[6] Examining positive affect towards Ultra-Orthodox children, which were not mentioned in the intervention, we find a positive albeit imprecisely estimated effect ($p = .18$). We interpret this pattern to suggest that our intervention affected students attitudes towards stigmatized groups mentioned in class, and may have had a more subtle, albeit imprecisely estimated impact on prejudice towards other outgroups which were not mentioned in class.

[6]We further discuss such ceiling effects in Appendix S7.

We further demonstrate that our intervention increased students' perceptions of similarity with multiple outgroups. The average treatment effect on our similarity index, reported on the right hand panel of Figure 3, is over 36% of an SD. This resembles approximately a half point shift on a five point similarity scale. When focusing on group-specific measures, we find that exposure to treatment increased perceptions of similarity with all outgroups, but effects on perceptions of similarity with Ultra-Orthodox children, which were not discussed in the intervention, are smaller and imprecisely estimated ($p = 0.22$).

In Figure 4, we further consider the effects of our intervention on four additional outcomes. First, we show that our treatment affected students behavior, increasing their registration for a future intergroup contact initiative by 17% of a SD, resembling a 5% increase in probability of registration amongst treated students. These findings are in line with effects on the contact intention index, which amount to a magnitude of 24% of a SD increase in response to treatment. When considering the intervention's effect on students' appreciation of diversity, the point estimate is positive but imprecisely estimated ($p = 0.19$). Finally we do not find any evidence that exposure to the intervention increased curiosity about outgroups, measured through a behavioral measure asking students to recommend additional groups to be featured in future seasons of the TV series "You Can't Ask That."

We subject our results to several diagnostic and robustness checks. First, we report balance on pre-treatment covariates in Appendix S4. Second, in Appendix S5, we demonstrate that minimal levels of attrition in the endline survey are not correlated with treatment status. Third, in Appendix S6 we consider the robustness of our results to a host of modeling specifications. Across all specifications, results remain consistent.

Taken together, we interpret the evidence in Figures 3-4 to suggest that participation in the intervention improved intergroup attitudes 7-14 days post-treatment. Specifically, we identify the most pronounced effects when focusing on groups mentioned in the intervention, and outcomes relating to willingness to engage with outgroups in the future. However, one might wonder whether treatment effects are the result of attitudinal change amongst prejudicial or non-prejudicial stu-

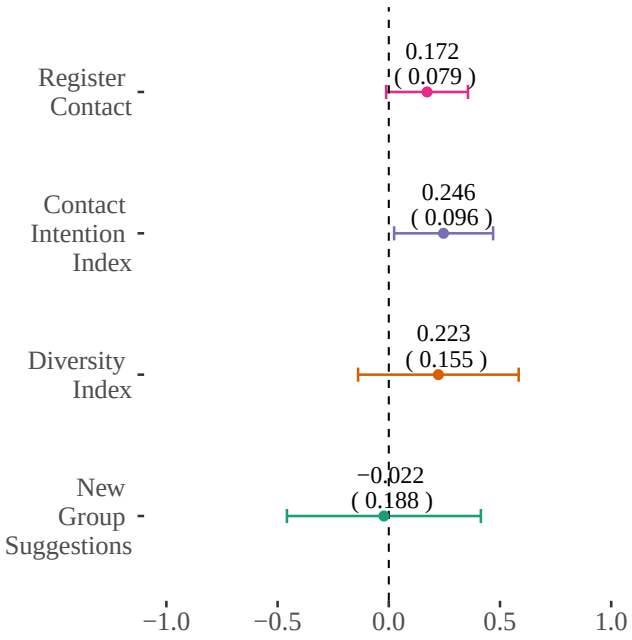

Figure 4: **Intervention Effects on Interest in Intergroup Contact, Appreciation of Diversity, and Curiosity about Outgroups.** Each point estimates and corresponding 95% confidence interval is extracted from an OLS model specified in Equation 1.

dents. In other words, was our intervention successful by preaching to the choir, or by generating attitudinal and behavioral change amongst the most prejudicial students. We consider this question in Figure 5, by plotting the distribution of our feeling thermometer index for treatment and control groups, pre- and post-treatment.

As reported in Figure 5, amongst control group students the pre- and post-treatment distributions of the feeling thermometer index are relatively similar. In contrast, for treated students, the pre- and post-treatment distributions of the feeling thermometer index are substantively different. Specifically, while the pre-treatment distribution of the feeling thermometer index peaks around 50 (blue line on the right panel of Figure 5), the post-treatment distribution peaks around 75 (red dotted line on the right panel of Figure 5). This difference is driven by post-treatment reduction in the concentration of students below and around the 50 score on the feeling thermometer index. We interpret these findings, as well as additional formal tests of moderation reported in Appendix S8, as evidence that our intervention was effective in large, by shaping the attitudes and behaviors of

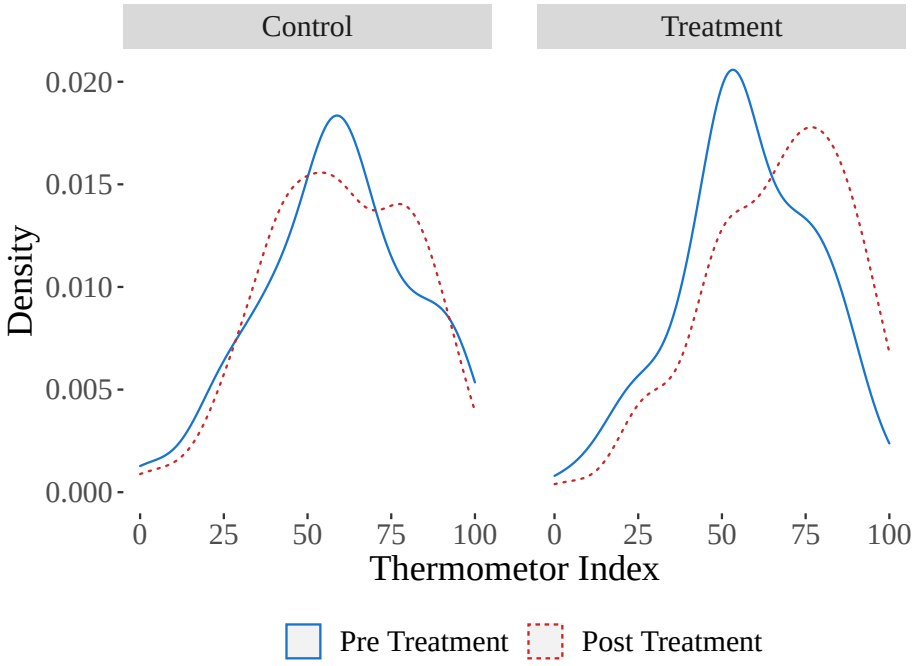

Figure 5: **Distribution Changes in Response to Treatment.** This plot visualizes the distribution of our feeling thermometer index amongst students from our treatment and control group, pre- and post-treatment.

prejudicial (rather than non-prejudicial) students.

Finally, since the implementation of our intervention coincided with a cycle of Jewish-Palestinian violence, we pay close attention to outcomes relating to Arab outgroups. In Appendix S9 we demonstrate that whereas attitudes towards Arabs improved between baseline and endline amongst treated subjects, similar attitudes were impaired amongst students in the control group. We cautiously attribute the negative trend amongst students in the control group to the cycle of violence which coincided with our intervention. In turn, our results emphasize how educational programs can be employed in times of intense intergroup conflict, to counteract the deterioration of intergroup attitudes and behaviors, and promote more favorable intergroup relations.

Discussion

These findings contribute to several theoretical and applied questions. First, we contribute to the literature on prejudice reduction, by laying out a novel and theoretically driven approach to reduce majority group members' prejudice towards minorities. We follow recent calls to adapt "elemen-

tal psychological forces in the service of prejudice reduction" (Paluck et al. 2021, p. 555), and demonstrate how sustained efforts to constructively broach prejudicial taboos, can be a productive pathway towards prejudice reduction. Our results emphasize that successful prejudice reduction interventions do not need to avoid sensitive topics at the heart of contentious intergroup relations, and that critical yet constructive confrontations of taboos can improve intergroup relations.

Second, we join a growing body of research which employs natural and field experiments in order to evaluate the efficacy of prejudice reduction interventions on attitudes and behaviors (Broockman and Kalla 2016; Hameiri et al. 2016; Goldenberg et al. 2018; Scacco and Warren 2018; Siegel and Badaan 2020; Mousa 2020; Lowe 2021; Weiss 2021). Following recent advances, we develop and test a theoretically motivated, intensive, and scalable intervention, in a naturalistic setting, measuring attitudes and behaviors, amongst our population of interest. In doing so, we depart from ongoing trends in the prejudice reduction literature which employ nudge-like interventions, and examine effects immediately post-treatment (Paluck et al. 2021). In turn, our findings emphasize how intensive and carefully curated educational programs can reduce prejudice, and serve to buffer the deterioration of intergroup attitudes in times of escalated conflict.

Finally, the timing of our intervention allows us to shed light on matters relating to temporal validity in the study of prejudice reduction. Many interventions to improve intergroup relations are implemented in divided societies *after* traumatic events of intergroup conflict (Paluck 2009; Scacco and Warren 2018; Mousa 2020). However, few studies focus on shaping attitudes and behaviors during cycles of intense intergroup conflict (Hameiri et al. 2016). This pattern is likely driven at least in part by the fact that common interventions such as intergroup contact initiatives (Scacco and Warren 2018; Mousa 2020), or canvassing efforts (Broockman and Kalla 2016), may not be well suited for times of intense conflict, during which members of different groups may be uninterested or threatened to engage with one another. Building on this insight, and acknowledging the direct threat that conflict poses to intergroup attitudes and behaviors, we designed our scalable intervention in a way that can be implemented even at times where direct intergroup engagement is not feasible. Our results emphasize that prejudice reduction interventions, that do not entail direct

intergroup contact, can serve as an important function in times of conflict, and may be employed in order to buffer deterioration of intergroup attitudes and behaviors.

Despite these contributions, our findings are not without limitations. First, like many field-experiments (Hameiri et al. 2016; Scacco and Warren 2018; Mousa 2020), the geographical scope of our research is somewhat limited. However, we emphasize that generalizability is rarely established through a single study, as it often entails cumulative efforts as part of a broad research program (Samii 2016). Though we find encouraging evidence when testing our intervention amongst our population of interest, we encourage future research to further investigate the generalizability of our results within and beyond Israel. Indeed, to facilitate cumulative learning, we designed our intervention in a way that can be replicated across-contexts, and tested at scale. To that end, we provide a full account of our intervention in Appenidx S2, and class materials in Appendix S11.

Second, though significantly improving on a majority of prejudice reduction interventions which measure outcomes immediately post-treatment (Paluck, Green and Green 2019; Paluck et al. 2021), our findings shed light on relatively short-term effects (1-2 weeks post-treatment). Regardless of duration, the effects we identify are substantively meaningful in our context. Indeed, they emphasize the importance of educational interventions in preventing the deterioration of intergroup attitudes and behaviors during cycles of intense conflict. That said, our theoretical framework, and the empirical context of classroom interventions, appear to be a suitable approach for explorations of longer-term effects (Dhar, Jain and Jayachandran 2022). We thus encourage future research to build on our theoretical framework and empirical design, and evaluate the durability of attitudinal and behavioral change, in response to interventions which broach prejudicial taboos.

Finally, though we elaborate on the theoretical framework underlying our intervention, and specify the theoretical mechanisms through which our intervention should work, our empirical focus is on evaluating the overall effect of the intervention, rather than identifying the relevant importance of each particular mechanism. Our empirical focus is driven by the notorious challenges of causally identifying mechanisms (Bullock, Green and Ha 2010), and by the understanding that effective prejudice reduction interventions may very well require "mixing of ingredients

from multiple theoretical perspectives" (Paluck et al. 2021, p. 555). Like recent landmark studies (Broockman and Kalla 2016), our intervention is likely effective through multiple mechanisms. Thus we encourage future research to employ lab, survey, and field experiments in order to build on our work, and further clarify the role of different mechanisms in generating the effects of our intervention.

References

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey Wooldridge. 2017. When should you adjust standard errors for clustering? Technical report National Bureau of Economic Research.

Allport, Gordon W. 1954. *The nature of prejudice*. Reading, Mass.: Addison-Wesley Pub. Co.

Brandt, Mark J. 2017. "Predicting ideological prejudice." *Psychological Science* 28(6):713–722.

Brauer, Markus and Abdelatif Er-Rafiy. 2011. "Increasing perceived variability reduces prejudice and discrimination." *Journal of Experimental Social Psychology* 47(5):871–881.

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

Bruneau, Emile G and Rebecca Saxe. 2012. "The power of being heard: The benefits of 'perspective-giving'in the context of intergroup conflict." *Journal of experimental social psychology* 48(4):855–866.

Bullock, John G, Donald P Green and Shang E Ha. 2010. "Yes, but what's the mechanism?(don't expect an easy answer)." *Journal of personality and social psychology* 98(4):550.

Choi, Donghyun Danny, Mathias Poertner and Nicholas Sambanis. 2021. "The Hijab Penalty: Feminist Backlash to Muslim Immigrants." *American Journal of Political Science* .

Dhar, Diva, Tarun Jain and Seema Jayachandran. 2022. Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. Technical report American Economic Review.

Er-rafiy, Abdelatif and Markus Brauer. 2013. "Modifying perceived variability: four laboratory and field experiments show the effectiveness of a ready-to-be-used prejudice intervention." *Journal of Applied Social Psychology* 43(4):840–853.

Falomir-Pichastor, Juan Manuel, Carmen Martínez and Consuelo Paterna. 2010. "Gender-role's attitude, perceived similarity, and sexual prejudice against gay men." *The Spanish Journal of Psychology* 13(2):841–848.

Fiske, Susan T. 1998. *Stereotyping, prejudice, and discrimination.* McGraw-Hill.

Goldenberg, Amit, Smadar Cohen-Chen, J Parker Goyer, Carol S Dweck, James J Gross and Eran Halperin. 2018. "Testing the impact and durability of a group malleability intervention in the context of the Israeli–Palestinian conflict." *Proceedings of the national academy of sciences* 115(4):696–701.

Hameiri, Boaz, Roni Porat, Daniel Bar-Tal and Eran Halperin. 2016. "Moderating attitudes in times of violence through paradoxical thinking intervention." *Proceedings of the National Academy of Sciences* 113(43):12105–12110.

Haslam, Nick and Steve Loughnan. 2014. "Dehumanization and infrahumanization." *Annual review of psychology* 65:399–423.

Hsieh, Wing, Nicholas Faulkner and Rebecca Wickes. 2021. "What reduces prejudice in the real world? A meta-analysis of prejudice reduction field experiments." *British Journal of Social Psychology* .

Kalla, Joshua and David Broockman. 2021. "Which narrative strategies durably reduce prejudice? Evidence from field and survey experiments supporting the efficacy of perspective-getting." *American Journal of Political Science* .

Kam, Cindy D and Camille D Burge. 2018. "Uncovering reactions to the racial resentment scale across the racial divide." *The Journal of Politics* 80(1):314–320.

Kinder, Donald R, Lynn M Sanders and Lynn M Sanders. 1996. *Divided by color: Racial politics and democratic ideals*. University of Chicago Press.

Kteily, Nour S and Kaylene J McClanahan. 2020. "Incorporating insights about intergroup power and dominance to help increase harmony and equality between groups in conflict." *Current opinion in psychology* 33:80–85.

Langer, Ellen J, Susan Fiske, Shelley E Taylor and Benzion Chanowitz. 1976. "Stigma, staring, and discomfort: A novel-stimulus hypothesis." *Journal of Experimental Social Psychology* 12(5):451–463.

Liberman, Zoe, Amanda L Woodward and Katherine D Kinzler. 2017. "The origins of social categorization." *Trends in cognitive sciences* 21(7):556–568.

Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7(1):295–318.

Lowe, Matt. 2021. "Types of contact: A field experiment on collaborative and adversarial caste integration." *American Economic Review* 111(6):1807–44.

MARBLE, WILLIAM, SALMA MOUSA, ALEXANDRA A SIEGEL et al. 2021. "Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes." *American Political Science Review* pp. 1–18.

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

Murrar, Sohad, Mitchell R Campbell and Markus Brauer. 2020. "Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap." *Nature Human Behaviour* 4(9):889–897.

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

Paluck, Elizabeth Levy, Seth A Green and Donald P Green. 2019. "The contact hypothesis reevaluated." *Behavioural Public Policy* 3(2):129–158.

Ramasubramanian, Srividya. 2011. "The impact of stereotypical versus counterstereotypical media exemplars on racial attitudes, causal attributions, and support for affirmative action." *Communication Research* 38(4):497–516.

Ramasubramanian, Srividya. 2015. "Using celebrity news stories to effectively reduce racial/ethnic prejudice." *Journal of Social Issues* 71(1):123–138.

Richards, Zoeë and Miles Hewstone. 2001. "Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change." *Personality and Social Psychology Review* 5(1):52–73.

Ron, Yiftach, Johanna Solomon, Eran Halperin and Tamar Saguy. 2017. "Willingness to engage in intergroup contact: A multilevel approach." *Peace and Conflict: Journal of Peace Psychology* 23(3):210.

Saguy, Tamar. 2018. "Downside of intergroup harmony? When reconciliation might backfire and what to do." *Policy Insights from the Behavioral and Brain Sciences* 5(1):75–81.

Saguy, Tamar, Nicole Tausch, John F Dovidio and Felicia Pratto. 2009. "The irony of harmony: Intergroup contact can produce false expectations for equality." *Psychological Science* 20(1):114–121.

Samii, Cyrus. 2016. "Causal empiricism in quantitative research." *The Journal of Politics* 78(3):941–955.

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.

Siegel, Alexandra A and Vivienne Badaan. 2020. "# No2Sectarianism: Experimental approaches to reducing sectarian hate speech online." *American Political Science Review* 114(3):837–855.

Smooha, Sammy. 2015. *Lo Shovrim et Ha-Kelim: Madad Yahasei Aravim-Yehudim Be-Yisrael 2015*. The University of Haifa and The Israeli Institute for Democracy.

Weber, Christopher R, Howard Lavine, Leonie Huddy and Christopher M Federico. 2014. "Placing racial stereotypes in context: Social desirability and the politics of racial hostility." *American Journal of Political Science* 58(1):63–78.

Weiss, Chagai M. 2021. "Diversity in health care institutions reduces Israeli patients' prejudice toward Arabs." *Proceedings of the National Academy of Sciences* 118(14).

Williamson, Scott, Claire L Adida, Adeline Lo, Melina R Platas, Lauren Prather and Seth H Werfel. 2021. "Family matters: How immigrant histories can promote inclusion." *American Political Science Review* 115(2):686–693.