Leveraging Virtual Contact and Social Networks to Foster Interethnic Harmony*

Abu Siddique† Michael Vlassopoulos‡ Yves Zenou§

December 2025

Abstract

This paper investigates whether virtual contact, initiated through a documentary film, can promote interethnic harmony. We carried out a cluster-randomized field experiment

*We thank Nathan Nunn, Larry Katz, four anonymous referees, Britta Augsburg, James Banks, Abigail Barr, Catia Batista, Sonia Bhalotra, Michael Callen, Maria Cubel, Amrita Dhillon, Paul Dutronc-Postel, Lucas Finamor, Leonie Gerhards, Shaun Hargreaves Heap, Eliana La Ferrara, Hikaru Kawarazaki, Cynthia Kinan, Sonya Krutikova, Gianmarco Leon-Ciliotta, Matt Lowe, Juanjuan Meng, Alice Mesnard, Chris Roth, Moses Shayo, Andrea Smurra, and seminar and conference participants at KCL, LSE, IFS, Oxford, Royal Holloway, Manchester, City, Erasmus, IFS/STICERD/UCL Development Seminar, Peking, Technical University of Munich, Birkbeck, Sussex, and NBER Political Economy fall meeting, CEPR-TCD-TIME Conference on Economic Development, ASSA 2024, Networks and Development Workshop, Monash Network Economics Conference, NOVAFRICA Conference, Bristol BAEM Development, the 2024 Discrimination and Diversity Workshop, Summer Workshop on the Economics of Inter-Group Relations and Integration, Patras Applied Microeconomics Workshop, and the CDES Sustainable Development Conference 2023 for constructive comments. We are very grateful to Patrick Healy and Ignasi Merediz Solà for their excellent research assistance throughout the project. We also thank Asif Ali Mahmud, Tarannum Azim Baigh, Nusaiba Binte Zakaria for their excellent research assistance. Ethical clearance for this project was received from the Monash University Human Research Ethics Committee (ref no. 35152) and is pre-registered at the AEA RCT Registry (AEARCTR-0010730). Asad Islam was named in the ethics application and the pre-registration but later chose to withdraw from the project. This project is funded by the Diligentia Foundation. Siddique also acknowledges funding from King's College London (his previous affiliation). We thank Kazi Robiul Alam and Golam Faruk Sarker from the University of Rajshahi, and Hopna Kisku from *Ashrai* NGO for their comments and guidance on the documentary film. We thank the production company *Chitrakkhi* for creating the documentary film, and the NGOs *SARCH* and *Ashrai* for their support in the fieldwork. This paper was previously circulated with the title "Leveraging Edutainment and Social Networks to Foster Interethnic Harmony".

†Department of Economics, Royal Holloway University of London; IFS; and IZA. Email: abu.siddique@rhul.ac.uk

‡Department of Economics, University of Southampton, and IZA. Email: m.vlassopoulos@soton.ac.uk

§Department of Economics, Monash University, CEPR, and IZA. Email: yves.zenou@monash.edu

involving over 3,300 households across 121 multiethnic villages in Bangladesh. We find that a documentary film, designed to humanize the ethnic minority Santals and evoke empathy among the ethnic majority Bengalis, increased the ethnic majority's prosociality toward minorities, though the strength of the evidence varies by treatment arm and outcome. Using emotion-detecting software to analyze facial expressions during the film viewing suggests that the documentary elicited emotional responses consistent with empathy. We do not find evidence that the intervention reduced the prevalence of negative stereotypes and discriminatory opinions toward minorities. In villages assigned to target network-central individuals, we find positive behavioral effects on untreated individuals, including Santals, and village-level administrative data suggest a reduction in police complaints in those villages. About five months after the intervention, we conducted a casual work field experiment involving 720 participants from the main intervention. In this task, pairs of ethnic majority and minority participants jointly produced paper bags for a local supplier under a piece-rate compensation scheme. We find positive treatment effects on productivity for both ethnic groups, with effects concentrated in villages where network-central individuals were treated. For the ethnic majority, increased prosociality, and for the ethnic minority, reciprocity or peer pressure may have contributed to increased productivity. Overall, our findings suggest that virtual contact and social networks may help promote harmony within multiethnic communities.

JEL: C93, D9, I31, J15, O12

Keywords: Ethnic discrimination, stereotypes, social network, virtual contact, emotions, prosociality, productivity, randomized experiment, Bangladesh.

1 Introduction

Interethnic tensions and prejudice are pervasive in many parts of the world with important implications for social cohesion and economic development (Alesina and La Ferrara, 2005). In such contexts, ethnic minority groups are often at the receiving end of exclusionary and discriminatory attitudes from the more powerful and dominant ethnic majority group, which severely impedes their socioeconomic progress. Overcoming these ethnic cleavages to ensure opportunities for the economic and social advancement of disadvantaged ethnic minority groups is a crucial policy challenge for many developing countries (Barron et al., 2023).

A relevant example of these relationships is found in multiethnic villages in northwestern Bangladesh, where the majority Bengali population coexists alongside the Santal ethnic minority. The Santals, one of the largest Indigenous communities in South Asia, have ex-

perienced a long history of marginalization and exclusion (Roy, 2012). Colonial-era policies dispossessed many Santals of their ancestral lands and subjected them to exploitative labor arrangements. More recently, they have continued to face economic disadvantages, including discrimination in agricultural markets and limited opportunities for social and economic mobility. These persistent inequalities have contributed to social distance and low levels of interethnic trust, often leading to tensions and conflict in multiethnic rural villages. Despite the deep roots of these tensions, even modest interventions that build understanding and empathy may still be effective in improving intergroup relations.

Against this background, our study evaluates a novel intervention designed to promote interethnic harmony within multiethnic villages by increasing awareness and understanding of the ethnic majority about ethnic minorities through targeted media exposure. Specifically, we conducted a cluster-randomized field experiment in 121 villages in northwestern Bangladesh to assess the impact of a documentary film portraying Santal lives and livelihoods. The film features three main layers of storytelling narrated by Santals that aim to (i) familiarize viewers with Santal culture, (ii) expose them to the economic and other social challenges faced by the Santals, and (iii) showcase examples of Santals who have excelled in education and their professional lives. Exposure to entertainment media (edutainment) has shown promise in changing norms and behavior in various domains through the channels of providing new information and changing preferences (La Ferrara, 2016; Grady et al., 2021).

In our context, viewing the film serves as a medium for virtual contact, which may influence the attitudes and behavior of the ethnic majority towards the minority outgroup, by increasing empathy and identification. We thus posit that the documentary could have an affective impact on viewers, influencing their attitudes and stance toward the outgroup (Petty et al., 2003; Lerner et al., 2015). Direct contact is often considered an effective way to reduce prejudice and stereotypes– provided certain conditions are met– by promoting understanding of the outgroup and fostering positive emotions, such as empathy (Allport, 1954; Paluck et al., 2019; Lowe, 2021). However, in our context, encouraging direct contact might not be as effective, given the unequal status of the two groups and the possible anxiety that interactions might induce, which could limit engagement and potentially lead to backlash.[1] The content of the documentary is particularly important in this regard: it does not seek to disrupt the socio-

[1]This challenge is not unique to our setting. Turner et al. (2007) note that religious and ethnic groups often remain largely isolated from one another, so direct contact can be difficult to establish.

economic hierarchy of the villages but rather aims to humanize the Santals and familiarize viewers with their historical hardships and relatable aspirations. Importantly, recent evidence highlights the unique power of storytelling in influencing people, as stories tend to resonate longer and more deeply than statistics alone (Graeber et al., 2024). Therefore, the indirect form of contact with the outgroup, facilitated by the documentary and presented with an entertaining veneer, emerges as a low-cost (approximately $13 per participant), scalable, and logistically feasible approach that can be implemented in contexts where direct contact may be socially difficult or impractical.

A second objective of this study is to investigate whether the intervention has larger effects on untreated individuals when network-central participants are targeted. This allows us to test whether broader behavioral impacts vary depending on who is treated, even though we cannot separately identify the mechanisms, such as diffusion, visibility, or perceived leadership, that might explain these differences. A growing body of literature demonstrates that social networks can be leveraged to enhance the diffusion of information and promote the adoption of behaviors in diverse areas such as technology adoption, microfinance, public health and education (Valente, 2012; Banerjee et al., 2013; Kim et al., 2015; Banerjee et al., 2019a; Breza and Chandrasekhar, 2019; Beaman et al., 2021; Alan and Kubilay, 2025). In our study, we leveraged the social relationships of the ethnic majority communities to target individuals with high diffusion centrality for exposure to the film in a subset of villages ("Central" villages), to investigate if the film's information would spread widely and shift the perceptions and behavior of those not directly targeted by the intervention. We therefore compare outcomes of untreated individuals in Central villages with those in the other study arms. Thus, to our knowledge, we provide the first empirical test of both levers–virtual contact and network targeting–in the context of ameliorating interethnic tensions.

We collaborated with two local NGOs in northwestern Bangladesh, Ashrai and SARCH, to carry out the screenings of a documentary film across three treatment arms.[2] In the first arm, we randomly selected ethnic majority Bengalis (RR) to watch the documentary (Treatment 'Random'). In the second arm, we included a mix of randomly selected Bengalis (RC) and individuals with high diffusion centrality (selected using the approach in Banerjee et al., 2019a) to watch the film (Treatment 'Central'). The final group served as a 'Control'

[2]Specifically, Ashrai coordinated the film's production and carried out the screenings, while SARCH was responsible for data collection.

and watched a placebo documentary. Individual screening sessions were arranged at each participant's home using tablets. Within each treatment arm, we also collected data from both the ethnic majority and minority populations who did not view the film, to measure the effects on untreated individuals. Estimates for RR and RC identify causal effects for randomly selected individuals; comparisons involving CC (the network-central individuals) are descriptive because centrality-based selection combines selection and treatment.

We evaluate the impact of the intervention by collecting data through five different methods: (1) Lab-in-the-field experiments to elicit prosociality (altruism and solidarity) toward the outgroup–as interaction with outgroup members has been linked to increased prosocial behavior towards them (Rao, 2019); (2) Administrative data on interethnic dispute complaints obtained from police stations and village counselors to corroborate if the prosocial behavior observed in the 'lab' translated into real-world behavior; (3) A casual-work field experiment to measure the productivity of multiethnic pairs of workers–as recent field experiments have documented the negative impact of diversity on the work productivity of teams in development contexts (Hjort, 2014; Afridi et al., 2020; Marx et al., 2021; Ghosh, 2025); (4) Photographs to capture facial expressions and infer emotions and empathy building during exposure to the documentary; (5) Quantitative and qualitative surveys to understand a variety of other behaviors, beliefs, and intentions among Bengalis, including interethnic trust (which we measured using an experimentally validated survey instrument by Falk et al., 2018), as well as several other potential channels.

Our first main result is that about three months after viewing the documentary film, treated ethnic majority Bengalis demonstrated more prosocial behavior towards ethnic minority Santals compared to their counterparts in the control group, though the strength of the evidence varies by treatment arm and outcome. Specifically, we find that randomly selected participants in the 'Central' treatment arm gave 14.7% more in the Dictator Game ($p < 0.01$) compared to the control group. In the Solidarity Game, their giving was 8% higher, though this effect is not statistically significant. In the 'Random' treatment arm, participants exhibited a 6.4% increase in giving in the Dictator Game, which is not statistically significant, but a significant 7.1% increase in the Solidarity Game ($p < 0.10$). Regarding trust toward ethnic minorities, treated ethnic majority participants in the 'Random' villages exhibited 11.8% greater trust, while randomly selected participants in the 'Central' villages exhibited 21.7% greater trust (both $p < 0.01$), compared to the control group. These results

suggest that the intervention may have encouraged some forms of prosocial behavior toward minorities, consistent with evidence from other interventions aiming to improve intergroup relations (Finseraas et al., 2019; Rao, 2019). However, we do not find significant effects on a range of self-reported measures of interethnic relations, including the number of interethnic friendships, social interactions with non-coethnics, and the amount charged to minorities for water.

To explore the potential mechanisms at play, we collected a rich set of qualitative and quantitative data through in-person interviews and surveys. Our second key finding is that around 82 percent of Bengalis in the treatment groups acquired new information and learned new things about Santals–primarily centered on the occupational struggles of the Santal people, their educational pursuits and aspirations, and their potential for economic success. We also find that Bengalis who watched the documentary expressed an increased willingness to help Santals. The support they intended to offer ranged from general assistance to economic aid (jobs, financial assistance, etc.). These intentions were primarily driven by humanitarian concerns and a desire to alleviate the suffering of the Santals. Additionally, Bengalis expressed a willingness to encourage their coethnic neighbors to help the Santals as well.

On the other hand, despite acquiring new information from the film, Bengalis do not significantly change their pre-existing biases and opinions regarding Santals, including the persistence of negative stereotypes and discriminatory opinions. The finding of changes in intergroup behavior without changes in attitudes is consistent with evidence from various interventions targeting prejudice reduction (Paluck et al., 2021), including a recent study focused on improving intergroup cohesion in India (Ghosh et al., forthcoming). We propose two potential explanations for this finding: (i) Cognitive vs. affective components (Tropp and Pettigrew, 2005; Turner et al., 2007)–changing cognitive aspects of prejudice (such as negative stereotypes) through indirect contact can be more challenging than altering affective components (such as emotions and feelings), which may be more responsive to storytelling approaches (Graeber et al., 2024). (ii) Resistance to change (Watson, 1971)–a single documentary film may not be sufficient to counter deeply ingrained generational biases, as Bengalis might resist modifying their views, even when presented with new information.

The documentary, however, induced an emotional reaction among the viewers. We analyzed the viewers' facial expressions by taking candid photos during the screening and used

emotion-detection software to detect the emotions that were triggered during the viewing.[3] Using this facial expression data, we find sadness to be significantly more prevalent ($p < 0.05$), particularly among network-central participants. This emotional response, potentially reflecting an empathic response to the film's content, may have increased engagement with the intervention. While we cannot precisely pin down the mechanisms, the observed positive shifts in prosocial behaviors towards Santals, such as altruism, solidarity, and trust, among both treated and untreated Bengalis are consistent with this interpretation.[4]

Our third main finding concerns effects on untreated ethnic majority Bengalis who were not part of the intervention. We find that untreated Bengalis in the 'Central' arm showed more altruism (20.9% higher than the 'Control' group, $p < 0.10$), solidarity (27.3% higher than the 'Control' group, $p < 0.05$), and trust (8.1% higher than the 'Control' group, $p < 0.05$) towards ethnic minorities. In contrast, we find no statistically significant effects in the 'Random' arm for these outcomes. Similar to treated Bengalis, we also find no evidence of changes in stereotypes toward minorities among untreated Bengalis. However, discriminatory opinions decrease in the Random treatment and increase in the Central treatment. Regarding effects on Santals, we find that those in both the 'Random' and 'Central' arms displayed greater trust towards Bengalis, increasing by 11% ($p < 0.05$) and 21.7% ($p < 0.01$) compared to the 'Control' group, respectively. However, we find no statistically significant effects on altruism or solidarity. In addition, in both arms, ethnic minority Santals reported higher subjective well-being than those in the 'Control' group ($p < 0.01$). We also find a weakly significant increase in food security among Santals in the 'Central' arm ($p < 0.10$), possibly reflecting increased support from the ethnic majorities. The more pronounced effects observed in the Central arm may reflect the fact that network-central individuals were treated in this arm. These individuals were selected for their high centrality, which makes them well-positioned to influence others within the village. They also exhibited stronger emotional responses to the film, which may have shaped how they engaged with the intervention. While we cannot determine whether these patterns reflect persuasion, visibility, credibility, or diffusion, the

[3] We obtained consent from participants about taking pictures at both baseline and before the screening. This was consistent across all arms, including the 'Control' group that watched a placebo film.

[4] This finding is in line with the literature on emotion and decision-making (Lerner et al., 2015), which demonstrates how emotions can heighten attention to a situation (Schwarz and Bless, 1991), activate goals (Zeelenberg et al., 2008), and facilitate information sharing (Berger and Milkman, 2012). Prosocial behavior can also serve an emotion-regulation function, including alleviating sadness or distress (Schaller and Cialdini, 1988). More broadly, empathy has been linked to greater prosociality (Sze et al., 2012), improved relationships with disadvantaged groups (Batson et al., 1997), and reductions in antisocial behavior (Okonofua et al., 2021).

results suggest that who is treated matters for broader community responses.

To corroborate these findings and investigate whether the improved post-intervention interactions translated into better village-level relationships, we analyze administrative data on dispute complaints sourced from local police stations and village counselors' offices. We find no significant change in complaints made to village counselors (which are made for arbitration or *shalish* purposes) across any treatment arm. However, when examining complaints filed at police stations involving more serious issues with substantial negative consequences for the parties involved, we find a significant reduction in Bengalis' complaints against Santals post-intervention ($p < 0.05$), but only in the 'Central' treatment arm. In contrast, there is no significant reduction in Santals' complaints against Bengalis ($p > 0.10$). This suggests that the reduction in formal complaints occurred specifically in villages where network-central Bengalis were among those targeted. While we cannot determine whether this pattern arose through behavior change among the targeted individuals themselves or through broader village-level effects, the evidence points to differences in impact depending on who is treated.

About five months after the documentary viewing, we returned to the study area and conducted a casual-work field experiment, in the spirit of Hjort (2014), involving 720 randomly selected participants from the two treatment arms and the control group of the main intervention. This experiment, which serves as our second endline, allows us to investigate whether our intervention improved workplace productivity in ethnically mixed teams. We recruited 360 Bengalis and 360 Santals from different villages and paired them to work for a local supplier of paper bags (locally known as *thongga*). It is important to note that all participating Bengalis from the treatment groups had previously watched the documentary film and none were *network-central*. Their task was to jointly produce paper bags, with one worker randomly assigned the role of a *preparer* (preparing the materials), while the other the role of a *finisher* (completing the product). This setup created a vertical interaction in the work environment. Roles were switched halfway through, allowing us to measure each worker's productivity both as a *preparer* and a *finisher*. Their joint productivity determined their earnings, which were split equally between the two and paid at a piece rate. Based on prior evidence of the negative impact of diversity on work productivity of teams (Hjort, 2014; Afridi et al., 2020; Marx et al., 2021; Ghosh, 2025), we expect that productivity may be influenced by one's attitudes toward their co-worker, and thus, through this channel, the intervention could potentially impact productivity.

We find a significant overall increase in productivity of approximately 5%, but only in the 'Central' treatment arm. Although the individuals in the productivity sample were not themselves central, the differences across treatment arms may still reflect broader effects of having network-central individuals originally targeted in those villages. Moreover, in this arm, both Bengalis and Santals significantly increased their productivity only in the role of finisher– the position most critical in determining their final earnings. What explains this rise in productivity? We interpret it as an expression of prosocial behavior toward co-workers, driven by an effort to increase their earnings (Rotemberg, 1994). Specifically, Bengali participants exposed to the documentary demonstrated greater prosociality toward Santals, motivating them to work harder to enhance their co-workers' income. In turn, Santals also raise their effort either to conform to the effort level of Bengali workers or because of perceived social pressure to do so (Kandel and Lazear, 1992; Mas and Moretti, 2009; Georganas et al., 2015). To support this interpretation, we examine *finisher* productivity separately, comparing those who started in this role versus those who became *finishers* upon swapping. For Bengalis, productivity is similar regardless of the order in which they assumed the two positions. On the other hand, for Santals, we find that the difference in productivity between the 'Central' and 'Control' groups is more pronounced when they worked as a finisher after swapping roles, implying some responsiveness to the higher productivity of the Bengalis when they served as the finisher first. These patterns align with our interpretation: elevated prosociality drives the productivity of Bengalis, while conformism and peer pressure drive the productivity of Santals. We present a simple model to illustrate these interpretations of our findings.

To summarize, our study sets out to address two main questions: (i) Does virtual contact contribute to promoting interethnic harmony? (ii) Does the intervention have larger effects on untreated individuals when network-central individuals are targeted? Our evidence, gathered through incentivized measures, self-reported beliefs and attitudes, complaints, and observations of behavior in a natural casual work setting, offers two key insights. First, while deep-seated beliefs may remain unchanged, behavior can be positively influenced–we detect positive impacts on prosociality, disputes, and team productivity. These findings are consistent with insights from the psychology literature regarding the stability of one's values and resistance to attitude change (Schwartz and Bilsky, 1990; Albarracin and Shavitt, 2018) and the difficulty in changing them through indirect contact (Tropp and Pettigrew, 2005; Turner et al., 2007). Second, the consistent and robust impacts of the intervention observed

in the treatment arm where central individuals were exposed to the intervention suggest that network targeting plays a key role for broader community responses.

Literature. This paper contributes to a recent literature in economics that studies the impact of media exposure, showing that it can be effective in altering attitudes and behavior in a wide range of domains (Jensen and Oster, 2009; Gentzkow and Shapiro, 2004; La Ferrara et al., 2012; Kearney and Levine, 2015). More closely related to our focus, several papers have examined the impact of media on interethnic conflict, often exploiting the expansion of access to television or natural variation in radio coverage to identify the effect of media. The prevailing evidence suggests that media tend to reinforce racial and ethnic stereotypes and exacerbate tensions (Yanagizawa-Drott, 2014; Wang, 2021; Ang, 2023). However, in contrast to much of the literature, which focuses on the negative impacts of media, our study contributes by demonstrating that media can also foster empathy, enhance perspective-taking, and improve interethnic relations. Similar to our study, a few others have shown evidence that media can also improve interethnic attitudes and trust (Banerjee and Datta Gupta, 2015; Blouin and Mukand, 2019).

Only a handful of previous studies have studied the effectiveness of edutainment interventions experimentally in developing-country settings (e.g., Tanguy et al. (2014); Berg and Zia (2017); Banerjee et al. (2019b); Green et al. (2019); Bjorvatn et al. (2020); Riley (2024)). Our contribution is to experimentally examine the impact of edutainment on improving interethnic relationships in a context characterized by a deep-rooted legacy of discrimination and marginalization against the minority group, a topic that has received limited attention.[5] We demonstrate that indirect contact with outgroup members, facilitated by an entertaining documentary, allows for the acquisition of new information about the outgroup. In this manner, the paper contributes to the literature showing that contact is an effective means to eliminate prejudice and stereotypes (Allport, 1954; Paluck et al., 2019).[6] Methodologically, employing a randomized experiment enables us to offer clean identification of a possible link between exposure to edutainment and changes in attitudes and behavior among the ethnically

[5] A few recent studies in political science and psychology have investigated edutainment interventions (Paluck, 2009; Murrar and Brauer, 2018; Weiss et al., 2023) and narratives (Audette et al., 2020) as a means to reduce prejudice toward outgroups. However, these studies have not addressed our specific context or the wide range of outcomes we examine. In addition, our study advances the understanding of the underlying mechanisms involved.

[6] The effects of contact on intergroup outcomes has been demonstrated in various settings exploiting random assignment into groups (Scacco and Warren, 2018; Rao, 2019; Lowe, 2021; Boucher et al., 2021; Corno et al., 2022; Anderberg et al., 2024).

dominant group, as well as the underlying mechanisms. Furthermore, we provide suggestive evidence that emotions evoked through exposure to media may act as an important channel for treating members of an outgroup with more empathy. This aligns with recent evidence that interventions emphasizing perspective-taking can promote social cohesion in ethnically mixed settings (Alan et al., 2021).

This paper also connects with the literature on network targeting. A growing literature demonstrates that social networks can be leveraged to enhance the diffusion of information and promote the adoption of behaviors across various domains (Banerjee et al., 2013; Kim et al., 2015; Banerjee et al., 2019a; Beaman et al., 2021; Zárate, 2023; Alan and Kubilay, 2025). Our contribution lies in that we provide evidence that targeting connected individuals can prove more effective in improving relationships between distinct ethnic groups. In doing so, we extend the scope of the existing literature on behavior adoption by demonstrating the potential of network targeting to not only influence individual behaviors but also foster intergroup cohesion.

Finally, a large literature in economics and the social sciences more broadly has been concerned with uncovering the nature, roots, and consequences of ethnic and racial discrimination using laboratory experiments, field experiments, natural experiments, and non-experimental approaches. Altonji and Blank (1999), Charles and Guryan (2011), Lang and Lehmann (2012), and Neumark (2018), provide general overviews. We contribute to this literature by providing field-experimental evidence of a new approach to improving interethnic relations.

2 Context, intervention, and conceptual framework

2.1 The context

Bangladesh is a suitable setting to study interethnic intolerance due to its ethnically diverse population. The country is home to around 45 different ethnic minority groups, which are culturally, racially, ethnically, and linguistically different from the majority Bengali population. Historically, ethnic minorities in Bangladesh have been subjected to stigma, marginalization, and discrimination by the ethnic majority across a variety of domains, including healthcare, education, employment, justice, and political representation (Roy, 2012).

Our study takes place in the Northwestern region of Bangladesh, specifically in the Rajshahi and Naogaon districts, where a significant portion of the Santal population lives. The

Santals are the second-largest ethnic minority group in the country.

Historical background. The Santals are among the earliest settlers in the Bengal region and belong to the Austro-Asiatic group of pre-Aryan communities (Shamsuddoha and Jahan, 2018). During the British colonial period, they were subjected to exploitative labor systems that effectively turned them into bonded laborers. The burden of repaying loans became insurmountable, trapping generations of Santals in a cycle of servitude. British land reforms and legal policies during this period further marginalized the Santals, stripping them of land rights and deepening their economic and social dispossession (Shamsuddoha and Jahan, 2018).

In the mid-19th century, the Santals led a major rebellion against British authorities and landlords, known as the *Santal Hul*, but were ultimately defeated, resulting in mass starvation, disease, and the deaths of 10,000 Santals (Andersen, 2022). Following the partition of India in 1947, the Santals, who were largely concentrated in the newly formed East Pakistan (present-day Bangladesh), continued to experience dispossession and marginalization (Roy, 2012). The absence of formal land titles made it difficult for many minority communities to assert their legal rights, leaving them vulnerable to land grabbing and exploitation by the majority Bengali population (Bleie, 2005).

Everyday life and relationship between Bengalis and Santals. Like other ethnic minorities in Bangladesh, the Santals continue to struggle with difficult economic conditions, limited educational opportunities, and inadequate healthcare. In our baseline sample (Tables A1 and A2), Bengalis have slightly more schooling than Santals (6.03 vs 5.31 years), higher monthly personal income (9,069 vs 7,150 Taka), and are less likely to work in farming (56% vs 73%), while household size is similar (4.77 vs 4.76).

Predominantly landless farmers, many Santals work as sharecroppers or wage laborers on Bengali-owned land, while others engage in small-scale farming, fishing, and handicrafts. Economic interactions between Santals and Bengalis are frequent, particularly in agriculture and labor markets. In local markets, they sell agricultural produce and handmade goods, though they often receive lower prices than their Bengali counterparts (Siddique et al., 2023). Their lack of alternative skills or financial capital leaves them vulnerable to exploitation by the ethnic majority, particularly during lean agricultural seasons (Faridul Islam, 2006). Moreover, nearly 80% of households in plainland ethnic minority communities (i.e., non–Hill

Tracts indigenous groups such as Santals, Oraon, Hajong, and Khasi; in our study area predominantly Santals) have no access to credit (The Daily Star, 2024). The average monthly per capita income is about 1,600 Taka, whereas the national average for rural areas was about 6,000 Taka in 2022 (The Daily Star, 2024).

Land ownership remains the most pressing concern for plainland ethnic minorities, such as Santals and Oraos, as they continue to face land grabbing and eviction by both the state and the ethnic majority (Macdonald, 2021). Many Santals report losing their land through legal loopholes, deceit, and outright force, leaving them in a precarious economic position. For instance, Santals were promised a share of forest resources during afforestation efforts in the 1980s but never received their due. Meanwhile, Bengali farmers regained control over their lands, while Santal lands remained occupied by the Forest Department, which continues to expand plantations (Faridul Islam, 2006).[7]

Even in public services, policies intended to support ethnic minorities are manipulated by influential ethnic majorities, leaving little to no actual benefits for minority communities (The Daily Star, 2016). Education remains an important barrier to the socioeconomic mobility of the Santal community. Cavallaro and Rahman (2009) argue that state neglect has resulted in poor educational opportunities for Santals, contributing to intergenerational poverty. In fact, about 70% of Santals have no formal education, a rate significantly higher compared to the illiteracy rates among Bengalis (Uddin, 2009). In addition, language barriers in education prevent Santal children from succeeding in Bengali-medium schools, leading to high dropout rates (Cavallaro and Rahman, 2009). These adverse experiences have led Santals to avoid competing with Bengalis (Siddique and Vlassopoulos, 2020).

Violence and other injustices. Interactions between Bengalis and ethnic minorities in Bangladesh often result in conflicts and acts of violence (Roy, 2012) that are attributed to deep-rooted cultural differences. Amin (2018) highlights that differences in cultural value orientations create sociocultural barriers, leading to the exclusion and disempowerment of ethnic minorities.[8] Religion is another important factor that has shaped intergroup relations in this context. Bengalis are predominantly Muslim, with about 8% Hindu, whereas Santals

[7] Regarding this issue, Macdonald (2021) quotes a Santal man: *"We were fooled and cheated out of our own lands. These days we don't have the proper ownership of our lands. I can't sell my land because my papers and documents aren't ready."*

[8] According to Amin (2018), Bengalis and Santals differ on four of five cultural value orientations: relation to nature/God, view of human nature, time orientation, and social relations. The only similarity is in activity orientation or the way of living.

primarily practice their traditional religion and worship a deity known as *Thakur* (Risley, 1891), although many also identify as Christian. While religious tensions are more pronounced between Muslims and Hindus, Santals tend to suppress their religious identities, avoiding conflict but also diminishing their cultural visibility (Uddin, 2009).

Violence against the Santals is widespread, often orchestrated by powerful Bengali elites. Land disputes are a central issue, with Bengalis using force to displace Santals from their ancestral land. For example, in a Rajshahi case, a local Bengali Member of Parliament colluded with the police to burn down Santali homes, forcing 2,500 families to flee the region (International Federation for Human Rights, 2017; Minority Rights Group, 2018). Years later, justice has yet to be served (Land Portal, 2022).

Bengali land grabbers are also known to hire armed groups to intimidate and forcibly displace Santals (Faiaz, 2021). Such attacks intensified during the COVID-19 pandemic when media presence was minimal, making it easier for perpetrators to act with impunity (Faiaz, 2021). More recently, instability following the 2024 “Students-People’s Uprising" in Bangladesh has led to an escalation of violence against minorities, including Santals (OWSA, 2024). This wave of attacks has left many Santals displaced and jobless and increased their vulnerability in an already precarious situation.

Women from ethnic minority groups in Bangladesh face a disproportionately high risk of rape and sexual violence, according to multiple rights reports, although official statistics are not disaggregated by ethnicity and thus do not allow per-capita comparisons (International Work Group for Indigenous Affairs, 2016; Amnesty International, 2016; Minority Rights Group International, 2024). Such crimes are rarely prosecuted, reinforcing a culture of impunity and deepening the insecurity experienced by ethnic minority women in Bangladesh.

Beyond direct physical violence, structural injustices also play an important role in the oppression of the Santals. Denial of access to irrigation water is a common form of economic coercion in which Bengali landowners deliberately block Santals from using water resources. This not only affects agricultural productivity but also pushes many Santal farmers into extreme despair. Reports indicate that such conditions have driven multiple Santals to suicide in recent years (The Business Standard, 2022; Partha, 2023).

2.2 Documentary film

Against the backdrop outlined in Section 2, we collaborated with the Bangladesh-based film production team, *Chitrakkhi*, operated by Bangladeshi film students, to produce a documentary film shedding light on the lives and livelihoods of the Santals in northwestern Bangladesh.[9] The film, called '*Ami Santal*' or 'I am Santal', features three main layers of storytelling: culture, economic conditions, and occupational success. See Table C1 for more details on the content of the documentary film. We focus on these three aspects based on studies in anthropology and sociology suggesting the prevalence of negative stereotypes among the ethnic majority regarding these aspects of the minority population in Bangladesh (Bal, 2007; Siraj and Bal, 2021). Firstly, the film familiarizes viewers with the Santali culture, showcasing aspects of their rituals, cuisine, etc., that are often unfamiliar to the ethnic majority. Most of this filming took place during the *Baha* festival. The focus then shifts to the trials and tribulations of the Santals, such as housing issues, lack of access to clean water, low incomes, labor market struggles, and the educational challenges faced by young Santals. Finally, viewers are introduced to the aspirations of the Santals, highlighting their potential for success, and showcasing how some Santals have overcome obstacles to achieve success in various domains, including education and profession.

All stories are narrated from a Santali perspective by local non-actor Santals. Filming took place outside the study region, in different villages from our study villages. Importantly, these stories depict the hardships Santals experience without attributing blame to Bengalis. The documentary aims to naturally capture the social issues and contexts without resorting to scripted narratives, serving as an ethnographic record of the lives of the Santals in Bangladesh. To maintain this authenticity, the film was created under the close supervision of visual anthropologist Kazi Robiul Alam and economic anthropologist Golam Faruk Sarker, both Professors of Anthropology at the University of Rajshahi, Bangladesh. We also received comments and suggestions on the film from Hopna Kisku, the Deputy Director of *Ashrai* (an NGO working for the welfare of ethnic minorities), who is also a Santal. The 45-minute documentary, illustrated in Figure A1, can be viewed at this link: YouTube.

To assess the documentary's emotive and moral sentiment, we analyzed its English transcript using LIWC-22 (Boyd et al., 2022). Full details of this analysis are reported in Ap-

[9]The director and editor, Labib Haque, has won numerous filmmaking awards and many of his films have been screened at international film festivals. A segment of this film was a finalist at the Dhaka OIC Youth Capital Film Competition.

pendix C. Table C2 shows that the film's tone is slightly negative-to-neutral (35.83 on a 1-99 scale), with minimal emotional language (2.17% negativity/sadness) and a high authenticity score (65.93), suggesting a genuine narrative. Prosociality-related language is nearly absent (0.65%), indicating that the documentary does not contain overt prosocial messaging. When compared to 74 Bangladeshi documentaries and a general film database (19,970 films), Figure C1 shows our documentary's emotional content is similar to other Bangladeshi documentaries, but lower than that of general films, indicating a less sensational tone. Figure C2 shows that while sadness-related content is slightly higher than in Bangladeshi documentaries, it remains well below the levels found in general films. Figure C3 further illustrates that moral judgments are much less frequent in our documentary than in other documentaries and general films, suggesting a more observational storytelling style that allows viewers to form their own interpretations without imposing strong moral evaluations.

2.3 Conceptual framework

We present here a simple conceptual framework to illustrate how the documentary might influence the prosocial behavior of the Bengali majority group toward the Santal minority group. In this framework, the majority group's prosocial actions are shaped not only by their own material interests but also by their concern for the well-being of the minority group.

We postulate that the impact of the documentary on viewers' willingness to help the minority operates through two channels: information about the minority group and empathy toward them. In turn, empathy is influenced by both the information obtained in the documentary and the viewers' baseline level of empathy.

Consequently, prosocial behavior is expected to be higher among treated Bengalis because the intervention enhances empathy, which increases the weight they assign to Santals' well-being. However, for information to increase prosociality, it is necessary for individuals to possess some baseline level of empathy; in other words, information and empathy act as complements. This simple framework highlights the channels through which the intervention might impact prosocial behavior, emphasizing the combined roles of emotional and informational components.

3 Research design and data

3.1 Research design

Sampling. We evaluate the effectiveness of the documentary film through a cluster-randomized field experiment in the Rajshahi and Naogaon districts of Bangladesh. We selected these districts because they contain large Santal settlements and our project partner, *Ashrai*, has operational coverage there. Santals (the second-largest ethnic minority group in Bangladesh) also live across other northwestern districts, so our choice reflects considerations of concentration and logistics, not exclusivity. In our sampled villages, the Santal share ranges from 13% to 83%, with the remainder being Bengalis.[10] Discrimination, conflicts, and violence between the ethnic majority (Bengalis) and the Santals are common in this region.

To facilitate the fieldwork, we partnered with two local NGOs, *Ashrai* and *SARCH*. From *Ashrai*, we obtained a list of approximately 150 multiethnic villages where the minority population constitutes between 10-85 percent.[11] That is, we considered villages where one group was the local majority and the other comprised at least 10% of residents, to ensure an adequate number of both ethnic groups per village and to facilitate more interaction between these groups following the intervention.

The study proceeded in the following sequence:

i. We randomly selected 121 villages from this list, and randomly assigned one-third to each of the three treatment arms (Control, Random, and Central). This randomization occurred prior to any data collection, so it did not include any stratification.[12]

ii. In the 41 villages assigned to the Central arm, we conducted a brief survey of approximately 18-20 random individuals per village, asking them to nominate 15 people they believed would be most effective at disseminating information. From these nominations, we selected the 7 highest-ranked individuals in each village.

iii. We then conducted a baseline survey across all 121 villages, aiming to select approximately 28 households per village.

[10] The largest ethnic minority group, the Chakma, resides in the Chittagong Hill Tracts region in the southeast, where traveling is restricted due to crime and safety issues.

[11] In the pre-registration, we specified an ethnic composition range of 30–70%, which was accurate based on older data from *Ashrai*. At baseline, more recent *Ashrai* data showed that selected villages ranged from 10–85%.

[12] Note that one additional village was mistakenly selected beyond what was pre-registered (120 villages), but we decided to retain it in the study to preserve statistical power.

iv. Within each village, the surveyed majority households were assigned to either watch the documentary assigned to their village (Santal documentary in Random and Central villages, flower-farming documentary in Control villages) or the untreated condition, while all minority households were assigned to the untreated condition. In Central villages, the seven pre-identified central households (CC) were excluded from the random treated draw, implying that treated "random" households in Central villages (RC) are mechanically less central than treated households in RR villages.

v. About three months after the baseline, the documentary screenings were delivered.

The study flowchart is given in Appendix Figure A2.

Enumerators were tasked with implementing a simple "skip" protocol in each village, visiting households door-to-door and ensuring they skipped one or two households after interviewing either the male or female head of each household. Since our study villages were new to the enumerators, we expected that the villages and households being surveyed would also be new to them, and that this procedure would approximate random sampling within villages. However, we cannot rule out deviations from the sampling protocol in practice, including the possibility of convenience or snowball-type recruitment once initial contacts were established.

Out of 3,500 participants initially approached, 3,347 (95.6%) completed the baseline survey, but we do not have systematic village-by-village counts of approached households or detailed enumerator routes, as this information was not recorded during fieldwork.[13] Comparing our sample to external benchmarks is also difficult, given the limited availability of disaggregated village-level statistics from these rural communities. We address the potential limitations of our sampling procedures through our empirical strategy, using post-double-selection LASSO for control variable selection, controlling for baseline outcomes when available, and reporting randomization inference p-values.

The male or female heads of households in our main study sample were informed that the NGO *Ashrai* would conduct an in-home documentary screening approximately three months later and that participation involved a baseline survey, the screening, and subsequent follow-ups, but they were not told the specific topic of the film. Participants were informed about the nature of the study, their right to decline participation or withdraw at any time, the

[13] While we do not observe individual or household characteristics for those who opted not to participate at baseline, one of the main reasons for non-participation was the absence of the household head.

use of photos, and the confidentiality of their responses. Consent was obtained verbally and in writing to accommodate varying levels of literacy among participants. For participating in the baseline survey, households were offered a bar of soap and 20 Taka ($1 = 100 Taka) top-ups on their mobile phones immediately after completing the survey. For participating in the screening, participants were informed that they could win up to 40 Taka in cash prizes by taking part in a short quiz based on the film being screened, and a chance to win a lottery of 5,000 Taka. At the endlines, participants had the opportunity to win cash prizes from decision-making experiments or the casual work experiment. Participants were told that the documentary would focus on people living in villages in the Rajshahi Division, and that screenings would take place privately in their homes using tablets.

The main sample and treatment arms. The treatment arms are as follows:

- Treatment *Random* (40 villages, $N = 562$ Bengalis): approximately 14 randomly selected ethnic majority households per village watched the documentary film. We refer to this group as 'Random in Random' or RR.
- Treatment *Central* (41 villages, $N = 562$ Bengalis): 7 ethnic majority households with members with high-diffusion centrality (referred to as 'Central in Central' or CC) and approximately 7 randomly selected ethnic majority households per village (referred to as 'Random in Central' or RC) watched the documentary film.
- Control (40 villages, $N = 568$ Bengalis): approximately 14 randomly selected households per village watched a documentary film on floriculture or flower farming in Rajshahi Division (of the same length as the documentary film on Santals).

Villages were randomly assigned to the Control, Random (R), or Central (C) arm prior to any baseline data collection to enable the centrality exercise in C villages. From each household, we targeted either the male or female head of the household for data collection and treatment delivery. In many households, additional members watched the film alongside the targeted respondent, so we treat exposure as occurring at the household level. We do not measure outcomes for non-targeted members and therefore cannot assess within-household spillovers.

The untreated sample. To examine the effects on individuals who were not directly exposed to the intervention, we also randomly selected participants in each treatment arm

who did not watch either the documentary film on Santals or the film on floriculture:

- Treatment *Random* (40 villages): $N = 270$ Bengalis and $N = 270$ Santals.
- Treatment *Central* (41 villages): $N = 280$ Bengalis and $N = 279$ Santals.
- Control (40 villages): $N = 277$ Bengalis and $N = 279$ Santals.

To provide more clarity on the sampling within each village, Appendix Figure A3 provides a visual representation of the village-level sample selection process. Similar to the main participants who watched the documentary films, the untreated sample also took part in both the baseline and endline data collections, and was offered the same incentives for taking part in surveys as the main sample. Out of the 3,500 participants initially approached, a total of 3,347 (95.6%) participated in the baseline survey.

Selection of individuals with high diffusion centrality. We identify network-central households in the Central arm using a brief nomination exercise, following the approach of Banerjee et al. (2019a). Prior to the baseline, we conducted a brief survey involving approximately 18-20 random individuals per Central arm village. We asked them to nominate the 15 people in their villages whom they believed would be the most effective in disseminating information. Our enumerators visited central locations in each village in the *Central* arm, such as the village market, where they randomly surveyed passers-by to create this list.[14] [15] From this list, we selected the seven most consistently and highly ranked names for the *Central* treatment.[16]

Screening. Screenings were conducted individually at each participant's house. To ensure that all screenings within a village were completed on the same day, simultaneous screenings took place across different households. Screenings ran between November 1, 2022 and January

[14] As noted by an anonymous referee, it is possible that the increased presence of the research team in the Central villages may have made the intervention more salient in these villages, potentially contributing to broader village-level responses. However, we believe this is unlikely to have had a significant impact on the results, as these villages on average were large (average size of 170 households), and the enumerators, who were local and spoke the same dialect as the participants, did not provide any information about the nature of the intervention during the initial survey.

[15] The specific question we asked was: *"If we want to spread information to everyone in the village about events, immunization programs, new loan projects, or a fair that we plan to organize, whom should we approach? Please nominate 15 such individuals/households from your village."*

[16] Note that we collected this information only in the 'Central' arm. We also do not know if the passers-by surveyed were residents of the same village or nearby ones. However, since we could successfully match most names with actual village residents, we believe those surveyed at village markets had a good understanding of the villages and their residents.

5, 2023. Using the mobile numbers collected at baseline (also used for participation top-ups), enumerators called ahead to schedule a day and time for each visit. The documentary was displayed on 8-inch tablets, allowing participants and other household members to watch and listen to the documentary together. At the end of the screening, there was a four-question quiz based on the film. Correctly answering each question allowed participants to win 10 Taka (chance to win a maximum of 40 Taka, approximately one-seventh of their daily income).[17] We organized this incentivized quiz to encourage participants to pay attention to the video content and reduce attrition. There was a break at the 25-minute mark of the video.

3.2 Data collection

We collected data at baseline and at two endlines. While the baseline data collection only involved surveys, the endline data collection involved self-reported surveys and a lab-in-the-field experiment (first endline), and a field experiment (second endline). We collected data from approximately 28 participants per village (one per household), involving both the main and untreated samples. Specifically, around 14 of these participants watched the film and are considered part of the 'main' sample, while the remaining 14 did not watch the film and constitute the untreated (non-viewer) sample, resulting in a total of 3,347 participants.

The baseline data collection was conducted in September-October 2022. The first endline took place in February-March 2023, while the second endline in April-May 2023. Since the screenings were conducted individually, we scheduled the first endline approximately 2.5-3 months after the screening for each participant, and the second endline (casual work experiment) approximately 4.5-5 months after the screening. This ensured that the time gaps between the screening and the endlines were similar for all participants. Figure A2 summarizes the timeline, and Appendix D.3 provides additional logistical details.

Lastly, we also collected administrative data on interethnic complaints from both police stations and village counselors' offices at both baseline and endline. These records are available only at the village level, as names and addresses were not disclosed, preventing linkage to individual participants.

[17] Figure A5 shows the proportion of participants in each treatment group who answered the quiz questions correctly.

3.2.1 Outcomes

We categorize outcomes into three domains: (i) *Behavior:* incentivized prosocial choices and self-reported behaviors; (ii) *Beliefs & Attitudes:* stereotypes and discriminatory opinions; and (iii) *Other outcomes:* mental health, subjective well-being, intercultural competence, and (for ethnic minorities only) economic outcomes. Within each domain, we report Westfall–Young FWER-adjusted and randomization-inference (RI) p-values. Consistent with our design, comparisons for RR and RC identify causal effects for randomly selected individuals, whereas CC-Control comparisons are descriptive because CC participants were selected on centrality. Appendix Table A4 maps each pre-registered outcome to its domain/family and reporting location.

Behavior. At the first endline, our main outcomes of interest for both ethnic groups were altruism, solidarity, and trust towards non-coethnics. Altruism and solidarity were measured using incentivized lab-in-the-field experiments conducted on the entire sample, using established behavioral games: the Dictator game (Eckel and Grossman, 1996) and the Solidarity game (Selten and Ockenfels, 1998). In the Dictator game, a sender unilaterally allocates part of an endowment to a passive recipient. In the Solidarity game, players precommit how much they would transfer to an "unlucky" recipient who may randomly incur a loss (decided by a coin-toss), thereby capturing willingness to insure non-coethnics under shared risk. Participants received compensation for one randomly selected game. Due to logistical constraints associated with implementing the incentivized Trust game (Berg et al., 1995) sequentially on a large sample size of over 3,000 participants, we opted for the measure of interethnic trust proposed by Falk et al. (2018). This measure relies on the following experimentally-validated survey question: "I assume that Santals (Bengalis) have only the best intentions".[18] Self-reported behavior toward non-coethnics includes Expected solidarity, Friends (share with non-coethnics), Water price (charged to non-coethnics), and Help (to non-coethnics). Interactions & visits comprise Interactions with non-coethnics, Interactions with coethnics, Visits to non-coethnics, and Visits to coethnics. Detailed outcome definition and experimental

[18] For the Dictator and Solidarity games, participants received an endowment of 100 Taka (equivalent to $1) in 5 and 1 Taka coins, amounting to nineteen 5-Taka coins and five 1-Taka coins for each game. While the Dictator and Solidarity games allowed for a maximum of 24 coins to be allocated, we used a 0-20 scale for the trust question to create a broadly comparable range across our measures of altruism, solidarity, and trust (presented in that order). This made it easier to explain the scale of the trust question to our non-standard 'subjects'. We also used a thermometer scale to record giving and trust at the endline with decimal precision (i.e., in *poisha*).

instructions are provided in Appendix D.1 and Appendix D.2.

Beliefs & Attitudes. We measured stereotypes and discriminatory opinions at baseline and endline. The Stereotypes index comprises six items on culture, profession, potential for success, and education. The Discriminatory-opinions index comprises seven items related to culture, education, honesty, profession, and work relationships. Each question uses a 0–10 agreement scale, where 10 indicates complete agreement. For the exact wording of the questions, please refer to Appendix D.1.

Other outcomes. To capture broader impacts, we measured mental health using PHQ-4 (Kroenke et al., 2009), four subjective well-being items from the World Values Survey, and intercultural competence. For ethnic minorities only, we additionally measured household food insecurity, new employment, and income. All of these outcomes were pre-registered, and detailed definitions are provided in Appendix D.1.

Dispute complaints using administrative data. We pre-registered the use of intereth-nic dispute data based on police complaints as village-level outcomes—Bengalis' complaints against Santals, and Santals' complaints against Bengalis. These complaints, collected from two police stations (*thana*) and covering our 121 villages, were gathered at both baseline and endline, concurrently with the surveys. In addition, we obtained complaint data made to local village councilors from their respective offices. The complaints made to police stations typically involve more extreme and violent matters, whereas those made to village councilors primarily relate to arbitration, locally known as *shalish*. We use both sets of village-level count variables as outcomes. These data allow us to distinguish between complaints made by ethnic majorities against minorities and complaints made by minorities against majorities. However, it does not allow us to identify the names of the individuals who filed them or the relative severity of the complaints.

Work productivity using a field experiment. At the second endline, about 4.5 months after the intervention ended, we conducted a casual-work field experiment in partnership with a local supplier of paper bags. For this experiment, we recruited 720 participants from the main study—360 Bengalis and 360 Santals—equally selected from the three treatment arms. We randomly sampled households to invite and over-invited to ensure Bengali–Santal

pairing; on a work day, if more invitees arrived than available space, the first arriving invitees were enrolled, and later arrivals received a show-up fee. This field experiment involved two participants—one ethnic majority and one ethnic minority—engaging in a casual work opportunity that lasted for 3 hours. Workers in pairs were mostly drawn from different villages, and often from different union councils. The work was paid at a piece rate, and individual earnings depended on joint productivity. This field experiment and its outcomes were also pre-registered. Further details regarding the field experiment are provided in Section 6.

3.2.2 Potential mediators

Information and intention channels. After watching the documentary film and completing the incentivized quizzes, participants in the treatment arms were asked a series of open-ended questions by the enumerator. The first question was *"Name five new things you learned from the video today."* This open-ended question allows us to measure whether participants in the treatment arm received new information about the Santals, and if so, what kind of information. Note that the control group was not asked this question because they watched a documentary film on floriculture. While we had pre-registered that this question would be asked of the control group as well, it was ultimately not included during data collection.

The next questions that participants were asked were: *"Now imagine a Santal in your village is in similar conditions as the Santals in the film that you just watched... (2) What would you do in that situation? (3) Why? (4) What would you advise your neighbors to do in that situation?"* These three questions were designed to capture a more qualitative overview of participants' intentions and reactions. The responses to these questions can help in understanding the underlying mechanisms driving behavioral changes. Additionally, these questions were designed to encourage participants to deeply reflect on the social issues presented in the documentary and contemplate potential solutions. All questions were open-ended, primarily to minimize response biases.

Emotion and empathy channels. Given that films can affect emotions and help develop empathy that may have long-lasting impacts that influence behavior, we investigate potential *empathy* channels by measuring the emotions experienced while watching the documentary film. Field assistants took candid photographs (with consent) of each participant's face

at a random point during the screening.[19] Using the Emotimeter software (developed by reAImagine), which uses machine learning to detect emotions in portrait photos, we assess whether certain emotions, such as sadness, anger, disgust, happiness, and other emotions (seven types in total, summing to 100%), were triggered among each participant. This software assigns a score ranging between 0% and 100% to each emotion category, where a higher score indicates a stronger activation of that particular emotion (illustrated in Figure A6). Importantly, the scores for all seven emotion categories, including a neutral category, always sum to 100% This allows us to measure participants' emotions more objectively.

3.2.3 Baseline characteristics, outcomes, and balance

At baseline, we collected data on a wide range of individual, household, and village-level characteristics. These included but were not limited to factors such as age, gender, education, income, household size, occupation, village size, and village ethnic diversity. Individual and household-level data were gathered from all 3,347 households participating in the baseline survey. We obtained the two village-level characteristics from official records provided by our partner organization, *Ashrai*. Table A1 summarizes these characteristics along with the baseline values of all outcomes for the main (treated) sample, and Table A2 reports the same for the untreated samples (Panel A: Bengalis; Panel B: Santals), while also testing the differences between treatment arms. Participants are around 40 years old, predominantly male household heads, Muslim, and primarily engaged in farming. Note that participants in the Central arm, which includes the selected high diffusion-centrality individuals, are slightly more educated and have more personal income.

In both tables, difference p-values are obtained from regressions of each baseline characteristic or outcome on treatment-arm indicators with union council fixed effects and standard errors clustered at the village level (the unit of randomization). We find that some characteristics, such as education and income, and some baseline measures of prosocial behavior (for example, altruism and solidarity among untreated Santals), show statistically significant differences between treatment arms (education: RR–C $p = 0.01$, CC–C $p = 0.01$; income: CC–C $p < 0.01$, RC–C $p = 0.05$). In the main treated sample (Table A1), we report six pairwise comparisons among RR, RC, CC, and Control for each of 27 variables (27×6 =

[19] Because we do not know the exact time when each video began screening, we cannot match the timing of the photo with specific video content, so we cannot link particular scenes to measured emotional responses.

162 tests). Recall that comparisons involving CC are descriptive. For clarity, we group the difference p-values into randomized comparisons (RC–C, RR–C, RC–RR) and comparisons involving CC (CC–C, CC–RR, CC–RC). We also report omnibus joint tests of equality across arms. Focusing on the randomized comparisons that should be balanced by design (RC-C, RR-C, RC-RR), one of three omnibus tests rejects equality in the main treated sample (RR vs Control, $p < 0.01$). Among the comparisons involving the non-randomly selected Central individuals (CC-C, CC-RR, CC-RC), one of three rejects (CC vs RC, $p = 0.05$). Among untreated Bengalis (Table A2, Panel A), one of three omnibus comparisons rejects (Random vs Control, $p = 0.02$). Among untreated Santals (Panel B), one of three rejects (Random vs Control, $p < 0.01$).

Given the number of these balance tests, some low p-values are expected by chance. Nevertheless, the overall pattern may also be consistent with deviations from the intended household sampling protocol. We, therefore, treat potential sampling deviations as a potential limitation and address them empirically by using post-double-selection LASSO for control-variable selection, always including education, income, and the baseline outcome as unpenalized controls when available, and by reporting randomization-inference p-values alongside village-clustered standard errors in all main tables.

3.2.4 Attrition

Total attrition among the main Bengali sample at the endline is 9.7%. To investigate the possibility of differential attrition across treatment arms, we regressed an indicator for attrition (1 if missing at the endline or 0 otherwise) on treatment status, with union council (a rural administrative unit) fixed effects and standard errors clustered at the village level, as pre-registered. We find that attrition is 2 percentage points (pp) higher in *RR*, 1pp higher in *RC*, and 3pp lower in *CC* relative to the *Control* arm, though none of these differences are statistically significant ($p > 0.10$). Therefore, we find no evidence of differential attrition across treatment arms in the main sample.

In addition, 3% of the untreated Bengali sample and 13% of the untreated Santal sample attrited at the endline. Among the untreated Bengali sample, attrition is 2pp lower in the *Random* arm and 8pp higher in the *Central* arm relative to *Control*, but neither difference is statistically significant at conventional levels (both $p > 0.10$). However, among the untreated Santal sample, attrition in both the *Random* and *Central* arms is significantly lower than

in the *Control* (both $p < 0.05$), indicating differential attrition in this sample. While this pattern raises the possibility of bias in estimates for untreated Santals, we find that baseline characteristics remain largely balanced across treatment arms (see Appendix Table A2, Panel B), with only minor differences in two characteristics between the two treatment arms. Moreover, untreated Bengalis from three villages (all in the *Central* arm) and untreated Santals from nine villages (three in the *Central* arm and six in the *Control* arm) attrited entirely from the endline survey. Excluding these villages does not substantively affect our conclusions.

3.3 Hypotheses

1. The *Random* and *Central* arms will improve interethnic outcomes for treated ethnic majority (assessed using survey questions, lab-in-the-field experiments, a field experiment, and administrative data) compared to the *Control* arm.

2. The *Central* arm will generate larger positive effects on neighboring (untreated) ethnic majority participants relative to the *Random* arm.

3.4 Empirical method

Treatment effects. We estimate the effects on the main sample using the following regression specification:

$$Y_{1ijc} = \alpha + \beta_1 RR_{ijc} + \beta_2 RC_{ijc} + \beta_3 CC_{ijc} + \zeta Y_{0ijc} + \Gamma' X_{ijc} + \nu_c + \epsilon_{ijc}, \quad (1)$$

where Y_{1ijc} is the outcome of individual i living in village j, in union council c, measured at the endline. RR_{ijc} is an indicator for the 'Random' Treatment (where random Bengalis are targeted), while RC_{ijc} and CC_{ijc} are indicators for the randomly selected and influential Bengalis in the 'Central' Treatment, respectively.[20] Y_{0ijc} is the baseline analogue of the outcome. X is a vector of baseline controls that were selected using the post-double-selection LASSO procedure (Belloni et al., 2014); however, we always control for outcomes measured at baseline whenever they are available, and participants' education and personal income. ν are union council fixed effects, which we pre-registered to include (note that union councils are rural administrative units, and each union council consists of around nine villages). We

[20] When estimating effects on untreated individuals, we include indicators for Random and Central villages.

cluster standard errors at the village level, which is our unit of randomization. In addition, we report RI p-values in all main tables.

Following the Hypotheses outlined in Section 3.3, we expect that β_1, β_2, and β_3 will all be positive (but negative for stereotypes and discriminatory opinions); also, if untreated individuals in Central villages respond more strongly than those in Random villages, we might expect that $\beta_2 > \beta_1$. Note that care should be applied when interpreting β_3, as it reflects a comparison between 'central' individuals in the Central treatment and individuals in the control group and does not permit a causal interpretation due to non-random selection.

As research assistants visited ethnic majority households in person to screen the documentary film, participation was very high (96%).[21] Moreover, research assistants called households in advance (using mobile numbers collected at baseline) to schedule a convenient day and time for the screening visit, which reduced non-participation due to unavailability and increased the take-up rate.

Inference and multiple hypotheses testing. Since we test many hypotheses, we control the family-wise error rate using the Westfall–Young step-down procedure, computed within the pre-specified outcome families (1,000 resamples), and report these FWER p-values in the main tables (Westfall and Young, 1993). For *Behavior*, we define three families: Prosociality (Altruism, Solidarity, Trust), Self-reports (Expected solidarity, Friends, Water price, Help), and Interactions & visits (Interactions with non-coethnics, Interactions with coethnics, Visits to non-coethnics, Visits to coethnics). We also consider labor productivity and interethnic complaints as separate families.[22] For *Beliefs & Attitudes*, we treat Stereotypes and Discriminatory opinions as separate families, while *Other outcomes* constitute a single family. We also report randomization-inference (RI) p-values that respect the village-level assignment (1,000 replications), as pre-registered (Young, 2019).

Pre-registration. Our study is pre-registered with the AEA RCT Registry (AEARCTR-0010730). We list and explain all major deviations in Table A3. All other deviations from the pre-analysis plan, not reported in Table A3, are also documented throughout the paper,

[21] We do not have all the outcome data for the 4% of the sample that did not watch the documentary. Note also that participants in one village (in the *Random* arm) collectively decided to opt out of watching the film.

[22] When we take a more conservative approach by pooling all behavioral outcomes into one family (prosociality, self-reports, interactions/visits, and productivity) show very similar qualitative conclusions, with many effects (altruism, trust, and finisher productivity) remaining statistically significant among the *RC* treatment sample.

often in footnotes. Table A4 lists all pre-specified outcomes and indicates where they appear in the paper.

4 Results

4.1 Information acquisition

We begin by investigating whether the documentary delivered new information about Santals to the targeted Bengalis. To do this, we use data from post-screening interviews with open-ended questions, focusing on whether Bengalis acquired new information and, if so, the nature of this information. These interviews were conducted only with Bengalis in the 'Random' and 'Central' treatment groups, as those in 'Control' watched a documentary on flower farming unrelated to Santals' lives.

Immediately after finishing the documentary, we asked participants, *"Name five new things you learned from the video."* Field assistants recorded keywords based on responses, rather than full transcripts, to reduce the workload of field assistants.

We find that 82% of Bengalis who watched the documentary learned at least one new thing about Santals, with an average of 4 new pieces of information received. This learning rate was similar across the 'Random' and 'Central' arms. Figure A4 shows the distribution, with a median of 4 new pieces of information among those who learned something new. Importantly, over two-thirds gained four or more new pieces of information about Santals. Note that some Bengalis acquired even more than five new pieces of information, as enumerators were instructed to record all new lessons mentioned by the participants. Figures A7-A8 provide histograms with a breakdown of the most frequently mentioned new lessons/information that indicate that reported learning centered on Santals' economic circumstances (especially agriculture) and education. Overall, the post-screening interviews suggest that the documentary conveyed substantive new information to viewers.

4.2 Treatment effects on behavior, beliefs and attitudes, and other outcomes

Behavior We begin by discussing results on the three main behavioral outcomes collected during the first endline through lab-in-the-field experiments and an experimentally validated survey question: altruism, solidarity, and trust toward non-coethnics. Figure I presents the

average for each outcome by treatment, with the bars showing village-level clustered 95% CIs.

Following the intervention, raw means suggest that treated Bengalis in the randomized groups (RR and RC) displayed increased prosocial behavior towards Santals compared to the control group. We also observe elevated prosociality among network-central participants (CC), but these effects should not be interpreted causally. This pattern holds across all three dimensions of prosociality: altruism, solidarity, and trust. In terms of altruism, Bengalis in all three treatment groups (RR, RC, and CC) transfer significantly more resources to Santals than those in the control group (t-test: all $p < 0.10$). However, no significant differences emerged between the treatment groups themselves. Regarding solidarity, the CC group showed significantly higher transfers than both the control group (t-test: $p = 0.01$) and the RR group (t-test: $p = 0.03$). Finally, trust towards non-coethnics increased progressively across the groups, with Bengalis in the control group showing the lowest levels of trust, while those in the CC arm exhibited the highest (t-test: all $p < 0.01$, except for RR vs RC).

We next estimate treatment effects using specification 1. Table I presents results for all behavioral outcomes (columns 1-11), while Figure II summarizes treatment effects across domains. In Column 1, we observe that sharing with a minority in the dictator game is higher in the 'Central' treatment relative to the control group. This is true for both randomly selected and centrally selected individuals. In terms of magnitude, individuals in this treatment group give 5.9 and 6.5 additional Taka relative to the control group. Since the control group shares an average 40.5 Taka, the treatment effect amounts to about 15% and 16% additional sharing, respectively. Note that giving is also higher in the RR treatment, however, the effect is smaller and not statistically significant. Nevertheless, the tests reported at the bottom of the table indicate that we cannot reject the equality of the coefficients in any pairwise comparison. Under within-family FWER adjustment for Prosociality, altruism is significant in RC and CC (FWER < 0.01) but not in RR (FWER $= 0.11$); the corresponding RI p-values are 0.26 (RR), 0.01 (RC), and $<$0.01 (CC).

Turning to the solidarity game reported in Column 2, we observe again larger sharing by centrally selected individuals in the Central arm than the randomly selected individuals, though differences are not statistically significant ($p = 0.193$). Sharing in the Random arm also increases by about 7.1% ($p < 0.10$) relative to the control group. Again, we cannot reject pairwise equality of the coefficients in the Random arm and the two Central arms. RR and

CC remain significant after within-family adjustment (FWER = 0.04 and 0.001), whereas RC does not (FWER = 0.11); the corresponding RI p-values are 0.07 (RR), 0.20 (RC), and < 0.01 (CC).

Finally, in Column 3 we present results from the trust survey question. We find statistically significant improvements in interethnic trust in all three arms, with the largest gains in CC. Specifically, trust in CC improves by 54.6% relative to the control group. It also improves by 11.8% in RR and by 21.7% in RC, the difference in the treatment effects between these two arms being statistically significant ($p = 0.09$). Trust was measured differently at baseline and endline: at baseline, respondents answered verbally on a 0–10 numeric scale (mean 7.1), whereas at endline a 0–20 visual thermometer was used (mean 8.7, equivalent to 4.4 on a 0–10 scale). These formats are not directly comparable: visual thermometers can elicit more cautious or varied responses, and respondents may interpret them differently than numeric inputs. In addition, the film may have prompted respondents to answer more critically, lowering inflated baseline ratings. Importantly, the level shift is similar across control and treatment groups, so the apparent endline 'decline' likely reflects measurement rather than a true drop in trust. Finally, effects in all three arms remain robust to within-family adjustment (FWER < 0.01; RI < 0.01).

To address possible attrition bias, we conducted Lee (2009) bounds analysis on the prosociality outcomes; see Table A6. Compared to the unadjusted treatment effects on altruism, the lower bounds for the RR and RC groups are insignificant, suggesting that these effects are sensitive to attrition. However, the estimated bounds for the CC group remain positive and statistically significant. For the Solidarity and Trust outcomes, the conclusions remain unchanged.

It is important to note that treatment effects in the Random arm (RR) are generally smaller and less statistically robust than those observed in the Central arm. The stronger effects observed when network-central individuals are targeted may reflect their greater influence, stronger baseline prosociality, or other features of the Central arm design.

Beyond altruism, solidarity, and trust, Table I also reports effects on expected solidarity, minority friendships, the price charged for water, willingness to help, social interactions with non-coethnics and coethnics, and visit frequency (columns 4-11). Across the various outcomes considered, we observe that the intervention did not lead to significant changes in the three treatment groups, with the exception of the self-reported willingness to help (Col-

umn 7). Specifically, we find that individuals in the Random treatment (RR) and nominated individuals in the Central treatment (CC) express a higher willingness to help relative to the control group, an increase of about 6%. For RR, the effect remains significant after within-family FWER (FWER = 0.045), whereas for CC it does not (FWER = 0.270). This finding is consistent with the responses to post-screening qualitative interview questions on viewers' intentions to support and help ethnic minorities. We also find a marginally significant reduction in the amount charged to minorities for water in the Random treatment ($p < 0.10$), which does not survive FWER adjustment.

In addition, we do not observe any significant change in social interactions with either non-coethnics or coethnics (see columns 8–11 of Table I and Panel C of Figure II). Component items underlying the interactions index show no detectable changes for interactions with Santals or with Bengalis (Appendix Table A7). The observation that intergroup behavior shifts without corresponding changes in attitudes is in line with findings from Paluck et al. (2021) and Ghosh et al. (forthcoming).

To summarize, these findings indicate that behavior in the experimental games was different in the randomized treatment groups relative to those in the control group. Moreover, when comparing the behavior of randomly treated individuals across the two treatment arms (RR versus RC), it emerges that the most robust shifts are evident in the Central arm. One possible interpretation is that these differences reflect the broader village context in Central villages, where network-central individuals were among those treated. This suggests that targeting central individuals may have amplified village-level behavioral change. However, we cannot determine whether these patterns arise from behavior change among central participants themselves or from broader community responses.

Beliefs: stereotypes and discriminatory opinions We now turn to impacts on *Beliefs & Attitudes*. Table II reports stereotype and discriminatory-opinion indices (column 1) and components (other columns), in Panels A & B, respectively. Figure II (Panel B) summarizes the indices.

The stereotype questions were designed to align with the themes explicitly addressed in the documentary (see Table C1). For instance, the documentary visually depicted the interiors of Santal homes and aspects of their daily lives, which challenged the stereotype that "Santals are often unclean/unhygienic." Furthermore, it highlighted professional and educa-

tional aspirations by featuring Santals who expressed hopes for their children's futures (e.g., becoming doctors or continuing education) and by showcasing successful Santal professionals. These elements directly correspond to stereotype questions about Santals' suitability for specific professions or their perceived tendency to remain in agricultural work.

Despite the documentary addressing these themes, we find little to no evidence that it shifted majority beliefs. At the index level, effects on both Stereotypes and Discriminatory opinions are statistically indistinguishable from zero across RR/RC/CC (Figure II, Panel B; Table II, column 1). This pattern is also evident in disaggregated measures of stereotypes in Columns 2-7 in Panel A of Table II and discriminatory opinions in Columns 2-8 in Panel B of Table II. At the item level, only one effect survives within-family Westfall–Young adjustment: in RR, endorsement of "Santals would not make good doctors" increases (FWER < 0.001), an adverse movement. One interpretation is that the documentary's portrayal of Santals inadvertently reinforced occupational stereotypes, even while fostering empathy.

Trust is measured in two ways: (i) agreement that "I assume that Santals have only the best intentions," elicited on a 0-20 thermometer scale (Table I, Column 3) and (ii) agreement that "one can easily trust a Santal" (Table II, Panel B, Column 8). In our data, RR/RC estimates are positive for the first measure, while effects on the second item are small (mixed sign) and not significant after within-family Westfall–Young adjustment. Because the two items differ in wording and survey context (a standalone intentions/trustworthiness item versus an item embedded in the discriminatory-opinions battery), we report both and interpret divergence as evidence of limited robustness across trust proxies.

Based on the realized sample sizes and observed standard errors for the RR treatment arm, the study could detect changes larger than approximately 0.21 SD for stereotypes and discriminatory opinions (in absolute terms) at the 95% confidence level. For the RC arm, the MDEs are slightly larger, ranging from approximately 0.26 SD for stereotypes to 0.31 SD for discriminatory opinions. These magnitudes imply that the index-level nulls are consistent with limited updating rather than low power.

Other outcomes Turning next to the "Other outcomes" domain: mental health (PHQ-4), subjective well-being (SWB), and intercultural competence (ICC), we find no systematic effects (see Table A5 and Figure II, Panel C). Point estimates are small in magnitude across arms: for mental health, coefficients range from -0.04 to -0.24 on a 0-12 scale (lower is better);

for SWB, 0.03 to 0.31 on a 0-40 scale; and for ICC, -0.03 to 0.03 on a 0-4 scale.

4.3 Potential mechanisms

We next examine possible channels for the treatment effects documented above. We consider three main channels. The first pertains to whether the documentary influenced participants' beliefs, attitudes, and interactions with the minorities. The second channel relates to the emotional and empathy imprint stirred by the contents of the documentary film. Here, we examine the emotional impact of the narrative and visual elements of the documentary, to assess whether the film may have elicited empathetic responses or emotional connections that subsequently influenced participants' actions. The third channel examines Bengali people's intentions and motivations to help Santals, using a rich set of qualitative data collected through in-person interviews conducted immediately after the film screening. Finally, we assess two alternative explanations: experimenter demand effects and social desirability bias.

Stereotypes, discriminatory attitudes, and interactions. As shown above (Table II and Panel B of Figure II), the intervention did not shift stereotypes or discriminatory opinions. To understand how preexisting attitudes (see Figure A9) might have influenced the documentary's impact, we conducted heterogeneity analyses by baseline measures of stereotypes (Tables A8 and A9), discriminatory opinions (Tables A10 and A11), and village ethnic composition (Tables A12 and A13). The analysis indicates that Bengalis with stronger negative stereotypes and discriminatory views about Santals were more likely to show solidarity with them. However, this tendency was less pronounced among treated participants, particularly with respect to discriminatory opinions.

We do not observe any consistent pattern regarding the heterogeneous impacts on updating biased beliefs by baseline biased beliefs or village ethnic composition (a proxy for village-level contact). However, an interesting pattern emerged regarding network-central Bengalis. In villages with a higher proportion of Santals, these central Bengalis exhibited lower prosocial tendencies towards Santals. This pattern in behavior is similar to the findings in Siddique and Vlassopoulos (2020), which suggests that being a numerical minority can make Bengalis' ethnic identity more salient and potentially lead to harsher behavior towards Santals. Therefore, the main takeaway is that while the documentary film's impact may have varied slightly due to pre-existing biased beliefs and village diversity, it did not directly

change Bengalis' underlying stereotypes or discriminatory views towards Santals. This null finding is still informative in highlighting the importance of considering other mechanisms, such as emotional responses or empathy, in explaining the observed shifts in prosociality measures.

We cannot definitively isolate the reasons why biases and opinions remained unchanged despite exposure to new information. One possibility is that cognitive components of prejudice, such as stereotypes and discriminatory opinions, are harder to shift than affective responses through a single, indirect-contact intervention (Tropp and Pettigrew, 2005; Turner et al., 2007). A second possibility is that a short documentary is insufficient to undo deeply ingrained views shaped by long-standing social and economic segregation (Watson, 1971). A third possibility could be that respondents updated privately but did not express these changes, for instance due to social conformity (Asch, 2003; Ushchev and Zenou, 2020; Boucher et al., 2024)—expressing fewer negative stereotypes could be perceived as deviating from the social norm, given that such attitudes towards ethnic minorities are commonly held by Bengalis. However, we view this as unlikely. Survey responses were elicited privately, limiting social image concerns. Moreover, if conformity were driving the null effects on beliefs, we would also expect it to shape outcomes that are more publicly observable, such as trust and altruism, or effort in joint production tasks, yet we observe positive effects there. Finally, conformity would likely dampen village-level responses among untreated individuals, whereas we detect such effects, further weakening this explanation.

Emotion and empathy channels. We next examine whether the documentary elicited emotional responses consistent with an empathy-based channel. Empathy is the capacity to connect emotionally with others by imagining and understanding their feelings (Davis, 1980). Prior research suggests that media exposure can help build empathy toward people depicted in the media, with effects that can be long-lasting (Herrera et al., 2018). Empathy, in turn, has been shown to affect prosocial behavior (Sze et al., 2012), improve relationships with disadvantaged groups (Batson et al., 1997), and even reduce crime (Okonofua et al., 2021).

While objectively measuring empathy is challenging, emotional activation, such as feelings of sadness, can serve as a pathway to building empathy toward others (Fernández-Abascal and Martín-Díaz, 2019). To capture viewers' emotional responses, we analyzed Bengali participants' facial expressions during the screening using emotion-detection software. Field

assistants took candid photos of participants at random times while they were watching the documentary.[23] We then used the Emotimeter App (developed by reAImagine), which uses machine learning with a pre-existing training dataset, to identify the primary emotions being activated during the screening. Psychology research suggests that facial expressions are informative about emotion and are difficult to fake for emotions such as sadness, anger, and fear(Ekman, 1993).

We were unable to analyze the facial expressions of a subset of participants due to non-viewing, image quality, or non-consent, with similar rates across treatment arms.. We report the raw differences in Figure III and regression estimates in Table A14. The first observation to make here is that Sadness is the emotion that was triggered the most in the control group, followed by Anger and Happiness. Relative to this baseline, we find that the CC group exhibits significantly more instances of Sadness ($p < 0.01$). The RC group also shows an increase in Sadness compared to the control group, though this effect is only marginally significant ($p < 0.10$). In contrast, the RR group shows a small but statistically significant unconditional increase (p=0.03), which is not significant with controls (Table A14). Moreover, pooling the RR and RC groups reveals a marginally significant positive effect on Sadness (see Table A15).[24] This suggests that the treated, non-central ethnic majority may have also experienced an emotional response to the documentary.[25] Moreover, Treatment CC displays more instances of Happiness than the control group, with this effect being significant at 5%.

These patterns are consistent with an empathy-based channel: the heightened sadness observed among the viewers in the CC arm may have made the minority group's struggles more salient and increased their determination to help. It may also have prompted discussion within viewers' social networks, potentially contributing to the positive shifts in altruism, solidarity, and trust observed in both our main sample and the effects discussed in Section 4.4. While we cannot isolate the precise channels through which these broader effects emerged,

[23] We obtained informed consent from participants twice: once before baseline and again right before screening. Almost all participants consented to having their photos taken, with a few exceptions declining due to religious reasons. All screening took place during the day, in participants' front yard (locally known as *uthon*), to ensure the lighting was sufficient to take clear photos.

[24] We also collected self-reported measures of sadness (along with feeling happy, anger, and disgust) from participants immediately after they finished watching the film. We find a positive but statistically insignificant correlation between the two measures of sadness. This may be because the facial expression data were collected during the film and reflect moment-to-moment reactions to specific content, whereas the self-reported measure captures an overall feeling of sadness after the screening.

[25] As a robustness check, because the emotion measures are right-skewed and positive, we estimate specifications using log-transformed outcomes. The results are qualitatively similar, and the Sadness estimates are more precisely estimated in the log specification.

the patterns suggest that central participants' emotional engagement may have played a role.

Another explanation could be that people who are more empathic in general would be more moved by the documentary film, and hence, would be more prosocial towards the Santals. To explore this interpretation, we measured three types of empathic tendencies of participants using the Interpersonal Reactivity Index (Davis, 1983) (see Appendix D.1 for definitions). We test for heterogeneity by interacting these baseline scores with treatment indicators (see Table A16). We find limited evidence of heterogeneity: higher personal distress predicts greater altruism toward Santals, while among central Bengalis, higher empathic concern is associated with less solidarity.

Qualitative evidence. While the acquisition of new information (presented in Section 4.1) through the documentary film might not have directly altered Bengali viewers' preexisting beliefs and attitudes, it had the potential to influence their intentions to help the Santals. To investigate these motivations and potential actions, we included additional qualitative questions immediately after they watched the film. We asked open-ended questions focused on their willingness to support Santals, the specific forms of support they would consider, their reasoning, and their intentions to encourage other Bengalis in their villages. Since these questions were only asked to the viewers who watched the documentary on Santals, we cannot claim the following qualitative evidence to be causal. We discuss this analysis in detail in Appendices E and F. This analysis was not pre-registered, but we pre-registered that viewers of the Santal documentary film will answer these questions following the film screening.

Among the 38 unique answers obtained from Bengalis, they generally expressed an intention to help Santals (about 55%) or focused specifically on providing financial aid (about 21%). Poverty and daily struggles were the most frequently cited reasons for wanting to offer support (over 50% of respondents). Furthermore, Bengalis who watched the documentary film expressed a desire to encourage their coethnic neighbors to do the same: about half indicated a willingness to encourage neighbors to support Santals generally, and roughly 30% specifically mentioned encouraging financial assistance.

Alternative mechanisms. *Experimenter demand effects:* In our intervention, participants in all three treatment arms received a form of intervention: the two treatment arms viewed our documentary, while the control arm viewed a placebo documentary on flowers—making it an active control. While experimenter demand effects are generally less concerning in

active control group designs (Haaland et al., 2023), to alleviate any remaining concerns, we nevertheless adopt an approach similar to Chopra et al. (2024). At the end of the endline survey, participants were asked the following open-ended question: *"If you had to guess, what would you say was the purpose of this study? You will only get one guess, and if you guess it correctly, you will get 50 Taka."* Enumerators categorized participants' responses into one of eight options, with only one being correct. The results, summarized in Figure A10, indicate that less than 3% of the sample accurately predicted the purpose of the study. Note that participants had a strong incentive to answer this question correctly, as the reward for doing so (50 Taka) was equivalent to one-sixth of their daily income.[26] We then estimate treatment effects on prosociality and beliefs after dropping the 3% sample who correctly guessed the study's purpose. The results, presented in Table A17 show that all our findings remain robust. We also take a more conservative approach and drop responses of "understanding Santals' lives in the villages"; Table A18 reports these results, showing that many of the treatment effects on prosociality remain significant.

Social desirability bias: Another related concern is social desirability bias, which might arise because of the thematic similarity between the Santal film contents and survey questions. This similarity might still induce some biases, especially among participants more inclined to give socially desirable responses to survey questions. To address this concern, we closely follow the approach in Dhar et al. (2022). Using the 13-point social desirability bias (SDB) scale, we carry out heterogeneity analyses. Table A19 presents the analysis for prosociality and Table A20 for beliefs. Our analysis reveals limited evidence of heterogeneity in altruism within the 'CC' group (limited to network-central people). Detailed definitions of the SDB scale and its components are provided in Appendix D.1.

4.4 Effects on untreated individuals within treated villages

We next examine the possible impacts of the intervention on untreated ethnic majority (Bengalis) and minority (Santals) residing in treated villages.

On the untreated ethnic majority. Figure A11 plots mean levels for the three prosocial outcomes by treatment. The pattern that emerges is one of larger effects on the prosocial behavior of people in Central villages, where half of the treated Bengalis were nominated as

[26]Participants were informed if they had guessed correctly or not after the casual work field experiment. Those who guessed correctly received their money through mobile money transfers.

network-central. This observation is confirmed by the regression analysis of treatment effects presented in Table III Panel A, which also permits quantification of the effects. In the Central arm, we find significant increases in giving in the solidarity game (27.3% of the average in the control group; $p < 0.05$), which remains significant after within-family adjustment for multiple testing (FWER = 0.01), and more moderate improvements in trust (about 8.1%; $p < 0.05$), not significant after FWER adjustment (FWER = 0.15). The impact on altruism is also sizeable in the Central arm (20.9% of the average in the control group) but less precisely estimated ($p < 0.10$) and insignificant after FWER adjustment (FWER = 0.15). On the other hand, in the Random treatment arm, we find no statistically significant effects on any of these three outcomes.

We acknowledge that these effects are quantitatively substantial. However, the Central treatment differs from the Random treatment not only in terms of who was targeted, but also in how the intervention may have been perceived, implemented, or internalized within villages. Accordingly, we interpret these results as evidence that the intervention had broader reach when network-central individuals were among the treated, while remaining agnostic about the precise channels.[27]

Within the self-reports family, most notable is the effect on friendships in the Central treatment, where we find a significant increase in the number of ethnic minority friends. Moreover, we observe a reduction in water charges in the Central arm, which is not robust to FWER. Expected solidarity shows no robust change.

On social interactions, we observe two interesting patterns: untreated Bengalis visited and interacted more with other Bengalis in the Central arm (FWER = 0.03), and visited Santal households in the Random arm, with this effect being marginally significant (FWER = 0.06). Visits to coethnics in the Central arm are not FWER-robust (FWER=0.11).

We also estimate the treatment effects on negative stereotypes and discriminatory opinions among untreated Bengalis (see Table III, Panel B). Analogous to the effects on the treated Bengalis, we find no evidence of changes in stereotypes among untreated Bengalis in either arm. However, we find that discriminatory opinions decrease in the Random treatment (not robust after adjustment) and increase in the Central treatment ($p < 0.01$).[28]

[27] We conducted various robustness checks to determine whether these effects were driven by outliers or varied based on individual or village-level characteristics. Our analysis confirmed that neither factor influenced the results.

[28] Within the stereotype index, several subcomponents (hygiene, success, work, and education) move in opposite directions and largely offset in the composite. Within the discriminatory-opinion index, the increase

Note also that while prosocial behaviors improved among untreated Bengalis in Central villages, discriminatory opinions increased in this arm. This divergence may be consistent with backlash from heightened minority salience in Central villages, and it suggests that incentivized prosocial behavior and stated discriminatory opinions need not move together.

Results on other outcomes are presented in Table A22. Effects on PHQ-4 and ICC are small and not statistically significant after adjustment.

What might explain the stronger impacts on untreated Bengalis observed in the Central treatment? A natural possibility is that treating network-central individuals amplifies village-level exposure to, and discussion of, the documentary. Several features of central individuals may contribute to these effects.

Firstly, as shown in Figure A12, central individuals demonstrated higher levels of altruism toward minorities at baseline. In addition, we have seen earlier that they experienced larger impacts on their prosocial behavior toward the minorities (Table I), and displayed stronger emotional reactions to the documentary (Table A14). Secondly, central individuals might possess greater persuasive abilities and broader influence within their communities. These differences in baseline characteristics and post-treatment responses suggest that when central individuals are treated, they may influence others more effectively. However, we cannot isolate whether these observed village-level shifts result from behavioral contagion, increased visibility of the intervention, or other mechanisms.

We also explore whether spillovers vary with untreated Bengalis' proximity to nominated network-central individuals (walking distance and reported monthly visits).[29] We find no consistent pattern across outcomes, and the distance measure is incomplete (roughly three-quarters missing), so we interpret these correlations cautiously (see Appendix Table A21).

On the untreated ethnic minority. We next turn attention to the possible effects of the intervention on untreated Santals.

Table IV presents estimates on the three main prosocial outcomes (see Figure A13 for raw averages). We find a significant increase in Trust in both the Random and the Central treatment arms, which remains significant after FWER (FWER < 0.01 in Random and in Central). This effect is more sizeable in the Central arm than in the Random arm, with

in Central villages is driven mainly by items on segregated schools, children's friendships with Santals, and trust.

[29] Proximity measures are available only for untreated Bengalis.

the two effects being statistically distinguishable ($p < 0.01$). For the other two outcomes, altruism and solidarity, impacts are positive for both treatments, and larger for the Central treatment, though not statistically significant.

For behavior outside the games, visiting both coethnics and non-coethnics increases sharply: visits to Santal households and to Bengali households rise in both arms and remain significant after FWER (e.g., Visit San: FWER < 0.01 in both arms; Visit Ben: FWER < 0.01 in Random and < 0.01 in Central). Interaction measures show mixed patterns (e.g., more interactions with Bengalis in Central, FWER $= 0.02$), suggesting that the visit margin accounts for much of the movement. Among economic outcomes, food security improves marginally in Central villages (about 28%), while new employment does not change (see Table A23. Finally, subjective well-being increases for untreated Santals, which is more pronounced in the Central arm in which this measure improves by about 51%. For the remaining outcomes, we do not find statistically significant effects.

To understand how these outcomes came about, note that results in columns 9 and 10 of Table IV suggest that Santals in treated villages were more likely to increase their interactions with both Santals and Bengalis by visiting them. These visits indicate improved cohesion between the two ethnic groups that could explain the rise in food security (received help from Bengalis) and the consequent rise in trust toward the majority.

Alternative explanations: In the previous section, we presented a test indicating that experimenter demand effects are unlikely to explain our findings. The impacts on the untreated samples, who were not directly exposed to the documentary and, hence, not subject to demand effects, provide further reassurance against these alternative explanations.

5 Impacts on village-level interethnic disputes

To assess whether the intervention generated wider impacts on interethnic relationships in treated villages, we gathered data on interethnic disputes. Due to confidentiality concerns, the data provided by police stations and village counselors are anonymous and aggregated at the village level. We consider two types of complaints, those directed to village counselors and those reported to the police. In this context, complaints filed with village counselors often serve arbitration purposes (locally known as *shalish*). This traditional approach helps avoid the time-consuming, costly, and complicated formal litigation process and is typically

governed by the country's Arbitration Act of 2001. Because the primary goal here is to settle disputes (e.g., for land acquisitions, threats of harm, verbal abuse, etc.), outcomes tend to be less severe for both parties.

Police complaints, on the other hand, can have much harsher consequences, including jail time, job loss, or social exclusion. The reasons for police complaints are often similar to those for disputes brought to village counselors (e.g., land acquisitions, threats of harm, verbal abuse). However, police complaints also include more serious issues, such as physical abuse (e.g., slapping, fighting, harassment), theft (e.g., looting houses, stealing livestock or agricultural machinery, stealing rickshaw-vans), arson (e.g., burning houses), property damage (e.g., cattle damaging or eating crops, poisoning fisheries), and so on. These complaints may also lead to more expensive litigation procedures if they proceed to court. Therefore, the economic consequences of police complaints are significantly more severe and long-lasting than those filed with village counselors. Importantly, many disputes never reach the formal complaint stage, indicating that the complaints data we have here captures particularly serious cases.

Table V presents treatment effects on interethnic dispute complaints at the village level, which we collected two months after the intervention ended. Columns 1 and 3 report monthly complaints made by ethnic minority Santals against ethnic majority Bengalis, while Columns 2 and 4 report monthly complaints made by ethnic majority Bengalis against ethnic minority Santals. In the regression analysis, we control for village size (the total number of households), and village ethnic diversity (the proportion of Santal households). We also control for the number of complaints at baseline (a month before the intervention started). Our results indicate a reduction in both types of complaints made by Bengalis in the 'Central' treatment arm, with the reduction being statistically significant in the case of police complaints only ($p < 0.05$). This reduction is substantive, amounting to a 41.5% decrease compared to the control group mean, though it does not remain significant under a conservative multiple hypothesis adjustment that considers all complaint outcomes as a single family ($p = 0.35$). For complaints filed by Santals, the coefficients are negative but not statistically significant at conventional levels.

The decrease in police complaints by Bengalis in the 'Central' treatment arm may reflect broader village-level shifts in behavior associated with the targeting of network-central individuals. This shift in attitude may have reduced disputes, encouraged informal conflict

resolution, or reduced complaints filed by Bengalis possibly due to increased awareness of the strong consequences their actions could have for the already socially disadvantaged Santals. However, we cannot identify whether the decline in complaints arose from changes in the behavior of the central participants themselves or from broader diffusion, visibility, or normative shifts. Still, the pattern is consistent with the idea that who is treated within a village can shape village-level outcomes.

On the other hand, the lack of a significant change in complaints filed by Santals to police stations might be due to the existing power difference between the two groups. Note that Santals filed significantly fewer complaints against Bengalis even in the absence of the intervention, as seen in the control group average. This lower complaint rate among Santals in the control arm suggests that the power imbalance might have discouraged them from making formal complaints, potentially leaving less room for further reduction in their complaint behavior in the treatment arms.

6 The casual work field experiment

6.1 The setting

The majority of participants in our intervention were agricultural day laborers or farmers. We conducted our field experiment on casual work immediately after a lean season (April-May), a period of reduced agricultural work and increased financial strain, making casual work opportunities highly attractive (Kaur et al., 2025).

Drawing inspiration from Hjort (2014), we designed a casual work experiment involving pairs of workers, each from a different ethnicity. Our primary focus was on exploring interethnic interactions within a labor-market context and making comparisons across treatment arms. To this end, we carefully matched each majority member with a minority member and avoided altering the ethnic composition of pairs due to concerns regarding statistical power and logistical constraints. Therefore, our field experiment allowed a Bengali and a Santal to collaboratively undertake a small-scale manufacturing job. This job involved producing paper packets, or *thongga*, for a local supplier. See Figure A14 for an illustration of the task and work arrangements. The task lasted three hours, including instructions, training, and a break.

In this task, participants assumed two roles: the *Preparer*, who cut and folded old newspa-

pers into packets, and the *Finisher*, who applied glue and ensured the quality of the packets. After an hour, they switched roles, so each participant experienced both roles. This setup implies that participants could start in the first round as the finisher (referred to as "First") or take on the role in the second round (referred to as "Second"). Both roles require similar skills and are equally integral to the completion of the overall task. In practice, this type of work is typically performed by a single person, but for the purposes of our experimental design, we divided it into two distinct roles.

Participants' productivity directly impacted their earnings, with 4 Taka (=4 cents) paid per packet, split equally between the two workers.[30] We measured productivity twice, during the break and at the end, providing data for both workers in both roles. Since, on average, they roughly produce 50 packets during the session, they make on average 200 Taka for the team, thus 100 Taka each. Each participant received a participation fee and a task completion fee, each amounting to 50 Taka. Considering the typical daily wage in this context is approximately 300 Taka, the average earnings (200 Taka) represent roughly two-thirds of a day's income.

We invited 900 Bengali and Santal participants (450 from each group) who had previously participated in our baseline. In total, we recruited 720 casual workers, or 360 pairs, evenly distributed across each treatment arm. We extended invitations beyond our capacity to guarantee the formation of 360 majority-minority pairs for the field experiment.[31] Workers not selected for participation received a show-up fee from the local supplier. All Santals in this field experiment came from the untreated sample, and all Bengalis came from the main (random) sample, excluding the network-central Bengalis. Appendix Tables A24-A25 report baseline balance for this subsample. Among Santals, a subset of covariates differs across arms; we therefore report clustered and RI p-values and control for pre-specified covariates selected through LASSO in regression analysis.

There is evidence suggesting that ethnic prejudices between workers can operate in workplace settings and influence team productivity, even if it results in costs for the discriminatory worker (Hjort, 2014; Hedegaard and Tyran, 2018; Afridi et al., 2020; Marx et al., 2021; Ghosh, 2025). Therefore, we expect that in this task, one's attitudes toward their paired co-worker would influence the effort they put forward, thereby negatively impacting the earnings of

[30] We pre-registered to pay 2 Taka per packet but increased it to 4 Taka per packet after consulting the paper bag supplier.

[31] Nine of these 360 participants (three in each arm) had not watched the documentary.

both individuals. Our primary interest lies in treatment differences in workers' productivity. We expect that if Bengalis who were exposed to the documentary harbored goodwill towards Santals, they would likely exhibit higher productivity in one or both roles relative to those in the control group. Note that workers could readily infer partner ethnicity from appearance/cultural cues and from accent during conversation, so partner identity was evident without additional disclosure.

6.2 Results

The overall output (number of finished packets) produced by pairs of workers showed a marked increase in the 'Central' treatment in comparison to the 'Control' group, as shown in Figure IV, Panel A (t-test; p-value=0.05). This gain in output is sizeable, amounting to a 5% increase. Pairwise comparisons between the other groups, 'Random' vs 'Control' and 'Random' vs 'Central', show no statistically significant differences.

Furthermore, an interesting pattern emerged in the work process. Finishers were consistently provided with sufficient items to process, indicating a bottleneck in this role, whereas Preparers demonstrated spare capacity. This is visually represented in Figure IV, Panel B, which shows output by position and ethnicity. Preparers produced on average about 36% more items than Finishers were able to process.[32]

When considering treatment effects for Finishers, whose productivity determined earnings, both ethnic majority and minority participants in the 'Central' treatment exhibited higher productivity than their counterparts in the 'Control' group, with a 21% improvement for the ethnic majority and a 15% improvement for the ethnic minority. For Preparers, we do not find significant treatment differences, suggesting that the overall treatment differences we see in Figure IV, Panel A are driven by Finishers. Regression analysis that includes control variables selected through post-double selection LASSO corroborates these findings (see Appendix Tables A26, A27, and A28).[33] We also examine the impact on the proportion of errors made on the final product in the casual work task (see Table A29) and find that

[32]However, in about one-fourth of pairs, the Finisher produced more packets than the Preparer supplied (an average of about 8 packets). In these instances, Finishers often assisted with the Preparer's task to maintain workflow and productivity. However, our data on Preparer output only includes the packets prepared by the assigned Preparer, not any additional contributions made by the Finisher. As a result, the recorded productivity for Preparers reflects only their individual effort, while Finishers' assistance in preparation is not captured in the data. All results based on Finisher output remain robust, and in many cases even stronger, when we exclude pairs in which the Finisher produced more packets than the Preparer.

[33]Heterogeneity analysis by baseline measures of prosociality, stereotypes, and discriminatory opinions reveals no significant patterns (results available upon request).

errors decrease significantly in the Central arm ($p < 0.05$, column 1). When disaggregated by ethnicity, we find a marginal reduction in errors by the ethnic minority ($p < 0.10$, column 3) in the Central arm only but not the majority ($p > 0.10$, column 2). This implies that ethnic minority workers were slightly more careful in performing the task in the Central arm.

Further insights can be obtained by examining the productivity of Finishers (by ethnicity) separately when they were first or second (see Figure IV, Panel C). In particular, for the ethnic majority, the similarity in productivity, regardless of the order in which they assumed the two positions, suggests that the disparity between the Central and Control groups is predominantly due to the treatment effect and not a response to the productivity of the ethnic minority as Finishers. Conversely, for the ethnic minority, we find that the difference between Central and Control is more pronounced when they acted as second Finishers, implying some responsiveness to the higher productivity of the Majority when they served as the first Finisher.

Figure A15 presents the productivity of Preparers by ethnicity and order. No discernible pattern emerges here, which is consistent with our earlier observation that the task of Preparers could be carried out faster, meaning they had excess capacity.

6.3 Discussion of potential channels

What explains the rise in productivity we observe in pairs involving treated Bengalis? Our interpretation is that working harder is an expression of prosociality toward one's co-worker in an attempt to raise their income (Rotemberg, 1994; Dur and Sol, 2010). That is, Bengali participants who were exposed to the documentary feel more prosocial toward Santals and increase their efforts in order to boost the income of their co-workers.[34] In turn, Santals respond by also raising their effort, either to conform or due to peer pressure (Kandel and Lazear, 1992; Mas and Moretti, 2009; Georganas et al., 2015). To provide empirical support for this interpretation, we examine the productivity of Finishers separately based on whether they were first or second to assume this role (see Figure IV, Panel C). Interestingly, for Bengalis, productivity is similar regardless of the order in which they assumed the two positions. On the other hand, for Santals, we find that the difference between Central and Control is more pronounced when they acted as second Finishers, implying some responsiveness to the

[34] This interpretation is supported by evidence reported in Table A30 indicating a marginally significant positive correlation between productivity as a finisher and altruism measured in the first endline ($p < 0.10$).

higher productivity of the Bengalis when they served as the first Finisher. These patterns are consistent with elevated prosociality acting as a driver for Bengalis and conformity acting as a driver for Santals. Our findings resonate with findings in Bhalotra et al. (2023), who find that intergroup contact improves coordination in mixed groups.

6.4 A simple model

To interpret the casual work results, we use a stylized two-player sequential effort model. In a mixed Bengali–Santal pair, output (packets produced) is the sum of the effort choices of the first and second finisher, $q = e_f^i + e_s^j$, and each worker earns the piece rate wq. The first finisher chooses effort first, while the second finisher observes this choice and then chooses effort.

Workers value own income and may also value the partner's income (altruism), captured by $\alpha^i \geq 0$. Effort is costly, and we allow for an additional dislike of effort mismatches, captured by $\theta^i \geq 0$ (interpretable as reciprocity or peer pressure). This structure delivers simple comparative statics: increases in altruism among Bengalis raise their effort directly, and stronger effort-alignment motives can propagate effort increases across partners in mixed teams. This provides a compact interpretation of the productivity patterns in Figure IV.

Appendix Section B presents the full setup and derivations.

7 Concluding remarks

This paper provides evidence of the potential of documentaries as a tool for fostering positive interethnic relations by enhancing emotional engagement and empathy through indirect contact with the outgroup. We find a positive impact on prosocial behavior, but no effect on beliefs regarding the minority outgroup. In villages where network-central individuals were targeted, the intervention led to broader changes in behavior across the village, including among Santals. In addition, we observe a reduction in ethnic dispute complaints by the ethnic majority, particularly at police stations, and an increase in productivity in a labor market setting, where multi-ethnic pairs engaged in team production.

It is noteworthy that, while the documentary did not explicitly correct negative stereotypes, it appears to have operated through affective responses consistent with empathy. This suggests that the intervention may operate not by challenging existing stereotypes but by

fostering a deeper understanding and emotional connection.

These findings suggest that virtual contact, facilitated through a documentary, can contribute to improved interethnic relationships. However, several important open questions remain for future research to address. First, it is important to understand the broader scalability of offering this type of documentary at a large scale through mass media, in which the audience composition might differ from our captive experimental sample, as viewers would self-select into watching the documentary. Second, the pattern of results, such as positive effects on prosociality but weaker to null effects on other self-reported behaviors, interactions, and beliefs, together with evidence of emotional activation during film exposure, suggests that the intervention operated primarily through affective rather than cognitive channels. Without changes in cognitive beliefs or sustained interethnic contact, the observed behavioral effects may attenuate over time as the film's emotional salience wanes. This limitation highlights that assessing the long-term persistence of these positive effects is essential to understand the true potential of virtual-contact interventions. Third, examining how media-based approaches compare with structured contact interventions offering similar informational content would provide valuable insights into scalable strategies for promoting social cohesion. Fourth, further research is needed to validate the robustness of the substantial effects on untreated individuals observed in Central villages and to better understand the mechanisms at play–whether they stem from the influence of treated individuals, their network position, or other village-level dynamics. Finally, replicating this study in settings with different levels of interethnic tension would help establish the broader applicability of our findings.

Data and Code Availability. All data and code necessary to reproduce the results in this article will be deposited in a public repository upon publication. The replication package will include raw data, code to construct analysis files, and code to reproduce all tables and figures, together with a README.

Supplementary Material

An Online Appendix for this article can be found at The Quarterly Journal of Economics online.

Data Availability

The data and code underlying this article is available in Siddique et al. (2025), in the Harvard Dataverse, https://doi.org/10.7910/DVN/WJAOOA.

References

Afridi, Farzana, Amrita Dhillon, Sherry Xin Li, and Swati Sharma, "Using social connections and financial incentives to solve coordination failure: A quasi-field experiment in India's manufacturing sector," *Journal of Development Economics*, 2020, *144*, 102445.

Alan, Sule and Elif Kubilay, "Empowering adolescents to transform schools: Lessons from a behavioral targeting," *American Economic Review*, 2025, *115* (2), 365–407.

_ , Ceren Baysan, Mert Gumren, and Elif Kubilay, "Building social cohesion in ethnically mixed schools: An intervention on perspective taking," *Quarterly Journal of Economics*, 2021, *136* (4), 2147–2194.

Albarracin, Dolores and Sharon Shavitt, "Attitudes and attitude change," *Annual Review of Psychology*, 2018, *69*, 299–327.

Alesina, Alberto and Eliana La Ferrara, "Ethnic diversity and economic performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

Allport, G. W., *The Nature of Prejudice*, Cambridge, MA: Addison-Wesley, 1954.

Altonji, Joseph G and Rebecca M Blank, "Race and gender in the labor market," *Handbook of Labor Economics (Editors O. Ashenfelter & D. Card)*, 1999, *3*, 3143–3259.

Amin, Md Ruhul, "Cultural Value Orientation and Inter-Ethnic Relation: A Case of Majority and Minority Ethnic Group in Bangladesh," *International Journal of Research in Sociology and Anthropology*, 2018, *4* (3), 57–66.

Amnesty International, "Bangladesh: Submission to the UN Committee on the Elimination of Discrimination against Women," `https://www.amnesty.org/en/wp-content/uploads/2021/05/ASA1348632016ENGLISH.pdf` 2016. Accessed: 2025-10-10.

Anderberg, Dan, Gordon B Dahl, Cristina Felfe, Helmut Rainer, and Thomas Siedler, "Diversity and Discrimination in the Classroom," 2024.

Andersen, Peter B, *The Santal Rebellion 1855–1856: The Call of Thakur*, Taylor & Francis, 2022.

Ang, Desmond, "The birth of a nation: Media and racial hate," *American Economic Review*, 2023, *113* (6), 1424–1460.

Asch, Solomon E, "Effects of group pressure upon the modification and distortion of judgments," in "Organizational influence processes," Routledge, 2003, pp. 295–303.

Audette, Nicole, Jeremy Horowitz, and Kristin Michelitch, "Personal narratives reduce negative attitudes toward refugees and immigrant outgroups: Evidence from Kenya," *Vanderbilt University Center for the Study of Democratic Institutions, Working Paper*, 2020, *1*.

Bal, Ellen, *They ask if we eat frogs: Garo ethnicity in Bangladesh*, Institute of Southeast Asian Studies, 2007.

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

_ , _ , _ , and _ , "Using gossips to spread information: Theory and evidence from two randomized controlled trials," *Review of Economic Studies*, 2019, *86* (6), 2453–2490.

_ , Eliana La Ferrara, and Victor H Orozco-Olvera, "The entertaining way to behavioral change: Fighting HIV with MTV," Technical Report, National Bureau of Economic Research 2019.

Banerjee, Ritwik and Nabanita Datta Gupta, "Awareness programs and change in taste-based caste prejudice," *PloS one*, 2015, *10* (4), e0118546.

Barron, Patrick, Louise Cord, José Cuesta, Sabina Espinoza, and Michael Woolcock, *Social sustainability in development: Meeting the challenges of the 21st century*, World Bank Publications, 2023.

Batson, C Daniel, Marina P Polycarpou, Eddie Harmon-Jones, Heidi J Imhoff, Erin C Mitchener, Lori L Bednar, Tricia R Klein, and Lori Highberger, "Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group?," *Journal of Personality and Social Psychology*, 1997, *72* (1), 105.

Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak, "Can network theory-based targeting increase technology adoption?," *American Economic Review*, 2021, *111* (6), 1918–1943.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 2014, *81* (2), 608–650.

Berg, Gunhild and Bilal Zia, "Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media," *Journal of the European Economic Association*, 2017, *15* (5), 1025–1055.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, reciprocity, and social history," *Games and Economic Behavior*, 1995, *10* (1), 122–142.

Berger, Jonah and Katherine L Milkman, "What makes online content viral?," *Journal of Marketing Research*, 2012, *49* (2), 192–205.

Bhalotra, Sonia, Irma Clots-Figueras, Lakshmi Iyer, and Joseph Vecci, "Leader identity and coordination," *Review of Economics and Statistics*, 2023, *105* (1), 175–189.

Bjorvatn, Kjetil, Alexander W Cappelen, Linda Helgesson Sekei, Erik Ø Sørensen, and Bertil Tungodden, "Teaching through television: Experimental evidence on entrepreneurship education in Tanzania," *Management Science*, 2020, *66* (6), 2308–2325.

Bleie, Tone, *Tribal peoples, nationalism, and the human rights challenge: The Adivasis of Bangladesh*, Dhaka: The University Press Limited, 2005.

Blouin, Arthur and Sharun W Mukand, "Erasing ethnicity? Propaganda, nation building, and identity in Rwanda," *Journal of Political Economy*, 2019, *127* (3), 1008–1062.

Bochkay, Khrystyna, Stephen V. Brown, Andrew J. Leone, and Jennifer Wu Tucker, "Textual analysis in accounting: What's next?," *Contemporary Accounting Research*, 2023, *40* (2), 765–805.

Boucher, Vincent, Michelle Rendall, Philip Ushchev, and Yves Zenou, "Toward a general theory of peer effects," *Econometrica*, 2024, *92* (2), 543–565.

_ , Semih Tumen, Michael Vlassopoulos, Jackline Wahba, and Yves Zenou, "Ethnic Mixing in Early Childhood: Evidence from a Randomized Field Experiment and a Structural Model," 2021. CEPR Discussion Paper No.15528.

Boyd, Ryan L, Ashwini Ashokkumar, Sarah Seraj, and James W Pennebaker, *The development and psychometric properties of LIWC-22*, Vol. 10, Austin, TX: University of Texas at Austin, 2022.

Breza, Emily and Arun G Chandrasekhar, "Social networks, reputation, and commitment: evidence from a savings monitors experiment," *Econometrica*, 2019, *87* (1), 175–216.

Cavallaro, Francesco and Tania Rahman, "The Santals of Bangladesh," *The Linguistics Journal*, 2009, *Special edition*, 192–220.

Charles, Kerwin Kofi and Jonathan Guryan, "Studying discrimination: Fundamental challenges and recent progress," *Annual Review of Economics*, 2011, *3* (1), 479–511.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val, "Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India," 2020.

Chopra, Felix, Ingar Haaland, and Christopher Roth, "The Demand for News: Accuracy Concerns versus Belief Confirmation Motives," *The Economic Journal*, 2024, *134* (661), 1806–1834.

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

Davis, Mark, "A multidimensional approach to individual differences in empathy," *JSAS Catalog of Selected Documents in Psychology*, 1980, *80* (10).

Davis, Mark H, "Measuring individual differences in empathy: Evidence for a multidimensional approach," *Journal of Personality and Social Psychology*, 1983, *44* (1), 113.

Dhar, Diva, Tarun Jain, and Seema Jayachandran, "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India," *American Economic Review*, 2022, *112* (3), 899–927.

Dur, Robert and Joeri Sol, "Social interaction, co-worker altruism, and incentives," *Games and Economic Behavior*, 2010, *69* (2), 293–301.

Eckel, Catherine C and Philip J Grossman, "Altruism in anonymous dictator games," *Games and Economic Behavior*, 1996, *16* (2), 181–191.

Ekman, Paul, "Facial expression and emotion.," *American Psychologist*, 1993, *48* (4), 384.

Faiaz, Zarif, "Attacks, land grabs leave Bangladesh's Indigenous groups on edge," `https://www.aljazeera.com/news/2021/7/30/bangladesh-indigenous-groups-chakma-khasi-santal-land-grab` 2021. Accessed: 2024-04-10.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global evidence on economic preferences," *Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

Fernández-Abascal, Enrique G and María Dolores Martín-Díaz, "Relations between dimensions of emotional intelligence, specific aspects of empathy, and non-verbal sensitivity," *Frontiers in Psychology*, 2019, *10*, 1066.

Ferrara, Eliana La, "Mass media and social change: Can we use television to fight poverty?," *Journal of the European Economic Association*, 2016, *14* (4), 791–827.

_ , Alberto Chong, and Suzanne Duryea, "Soap operas and fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 2012, *4* (4), 1–31.

Finseraas, Henning, Torbjørn Hanson, Åshild A Johnsen, Andreas Kotsadam, and Gaute Torsvik, "Trust, ethnic diversity, and personal contact: A field experiment," *Journal of Public Economics*, 2019, *173*, 72–84.

Gentzkow, Matthew A and Jesse M Shapiro, "Media, education and anti-Americanism in the Muslim world," *Journal of Economic Perspectives*, 2004, *18* (3), 117–133.

Georganas, Sotiris, Mirco Tonin, and Michael Vlassopoulos, "Peer pressure and productivity: The role of observing and being observed," *Journal of Economic Behavior & Organization*, 2015, *117*, 223–232.

Ghosh, Arkadev, "Religious divisions and production technology: Experimental evidence from India," *Journal of Political Economy*, 2025, *133* (10), 3249–3304.

_ , Prerna Kundu, Matt Lowe, and Gareth Nellis, "Creating cohesive communities: A youth camp experiment in india," forthcoming.

Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli, "ChatGPT outperforms crowd workers for text-annotation tasks," *Proceedings of the National Academy of Sciences*, 2023, *120* (30), e2305016120.

Grady, Christopher, Alice Iannantuoni, and Matthew S Winters, "Influencing the means but not the ends: The role of entertainment-education interventions in development," *World Development*, 2021, *138*, 105200.

Graeber, Thomas, Christopher Roth, and Florian Zimmermann, "Stories, statistics, and memory," *Quarterly Journal of Economics*, 2024, *139* (4), 2181–2225.

Green, DP, A Wilke, and J Cooper, "Countering violence against women at scale: a mass media experiment in rural Uganda. 2019," 2019.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart, "Designing information provision experiments," *Journal of Economic Literature*, 2023, *61* (1), 3–40.

Hedegaard, Morten Størling and Jean-Robert Tyran, "The price of prejudice," *American Economic Journal: Applied Economics*, 2018, *10* (1), 40–63.

Herrera, Fernanda, Jeremy Bailenson, Erika Weisz, Elise Ogle, and Jamil Zaki, "Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking," *PloS One*, 2018, *13* (10), e0204494.

Hjort, Jonas, "Ethnic divisions and production in firms," *Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

International Federation for Human Rights, "Bangladesh: Impunity reigns for human rights violations against Santal indigenous people," `https://www.fidh.org/en/region/asia/bangladesh/bangladesh-impunity-reigns-for-human-rights-violations-against-santal` 2017. Accessed: 2024-04-10.

International Work Group for Indigenous Affairs, "State of Indigenous Women and Girls in Bangladesh: Issues and Concerns at a Glance," `https://iwgia.org/images/publications/0753_Briefing_Paper_State_of_indigenous_omen_and_girls_in_Bangladesh_October_2016.pdf` 2016. Accessed: 2025-10-10.

Islam, M Faridul, "Ethnicity, ethnic conflict and discrimination against ethnic minorities of Bangladesh," *Journal of Ethnic Affairs*, 2006, *2*, 27–31.

Jensen, Robert and Emily Oster, "The power of TV: Cable television and women's status in India," *Quarterly Journal of Economics*, 2009, *124* (3), 1057–1094.

Kandel, Eugene and Edward P Lazear, "Peer pressure and partnerships," *Journal of Political Economy*, 1992, *100* (4), 801–817.

Kaur, Supreet, Sendhil Mullainathan, Suanna Oh, and Frank Schilbach, "Do financial concerns make workers less productive?," *Quarterly Journal of Economics*, 2025, *140* (1), 635–689.

Kearney, Melissa S and Phillip B Levine, "Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing," *American Economic Review*, 2015, *105* (12), 3597–3632.

Kim, David A, Alison R Hwong, Derek Stafford, D Alex Hughes, A James O'Malley, James H Fowler, and Nicholas A Christakis, "Social network targeting to maximise population behaviour change: A cluster randomised controlled trial," *The Lancet*, 2015, *386* (9989), 145–153.

Kroenke, Kurt, Robert L Spitzer, Janet BW Williams, and Bernd Löwe, "An ultra-brief screening scale for anxiety and depression: The PHQ–4," *Psychosomatics*, 2009, *50* (6), 613–621.

Land Portal, "Bangladesh's Santals continue struggle for ancestral land," https://landportal.org/news/2022/11/bangladeshs-santals-continue-struggle-ancestral-land 2022. Accessed: 2024-04-10.

Lang, Kevin and Jee-Yeon K Lehmann, "Racial discrimination in the labor market: Theory and empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Lerner, Jennifer S, Ye Li, Piercarlo Valdesolo, and Karim S Kassam, "Emotion and decision making," *Annual Review of Psychology*, 2015, *66*, 799–823.

Lowe, Matt, "Types of contact: A field experiment on collaborative and adversarial caste integration," *American Economic Review*, 2021, *111* (6), 1807–1844.

Macdonald, Geoffrey, "The Challenges Facing Plainland Ethnic Groups in Bangladesh: Land, Dignity, and Inclusion," https://www.iri.org/resources/new-bangladesh-report-examines-needs-of-plainland-ethnic-groups/ 2021. Accessed: 2024-11-21.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and team performance in a Kenyan organization," *Journal of Public Economics*, 2021, *197*, 104332.

Mas, Alexandre and Enrico Moretti, "Peers at work," *American Economic Review*, 2009, *99* (1), 112–145.

Minority Rights Group, "Adivasis in Bangladesh," https://minorityrights.org/communities/adivasis/ 2018. Accessed: 2021-09-17.

Minority Rights Group International, "Universal Periodic Review (UPR) of Bangladesh, 44th Session (Nov 2023): Factsheet on the Situation of Indigenous Peoples in Bangladesh," https://minorityrights.org/app/uploads/2024/01/upr44-bangladesh-mrg-2pager-final.pdf 2024. Accessed: 2025-10-10.

Murrar, Sohad and Markus Brauer, "Entertainment-education effectively reduces prejudice," *Group Processes & Intergroup Relations*, 2018, *21* (7), 1053–1077.

Neumark, David, "Experimental research on labor market discrimination," *Journal of Economic Literature*, 2018, *56* (3), 799–866.

Okonofua, Jason A, Kimia Saadatian, Joseph Ocampo, Michael Ruiz, and Perfecta Delgado Oxholm, "A scalable empathic supervision intervention to mitigate recidivism from probation and parole," *Proceedings of the National Academy of Sciences*, 2021, *118* (14), e2018036118.

OWSA, "Violence Against Minorities in Bangladesh Amid Political Chaos," `https://owsa.in/violence-against-minorities-in-bangladesh-amid-political-chaos/` 2024. Accessed: 2025-01-12.

Paluck, Elizabeth Levy, "Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda.," *Journal of Personality and Social Psychology*, 2009, *96* (3), 574.

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

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

Partha, Pavel, "Water injustice and Santal farmer's suicide," `https://www.newagebd.net/article/211491/water-injustice-and-santal-farmers-suicide` 2023. Accessed: 2025-01-30.

Petty, Richard E, Leandre R Fabrigar, and Duane T Wegener, "Emotional factors in attitudes and persuasion," *Handbook of Affective Sciences*, 2003, *752*, 772.

Rao, Gautam, "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools," *American Economic Review*, 2019, *109* (3), 774–809.

Riley, Emma, "Role models in movies: The impact of Queen of Katwe on students' educational attainment," *Review of Economics and Statistics*, 2024, *106*, 1–18.

Risley, Herbert Hope, *The tribes and castes of Bengal: Volume II*, Vol. 2, Printed at the Bengal secretariat Press, 1891.

Rotemberg, Julio J, "Human relations in the workplace," *Journal of Political Economy*, 1994, *102* (4), 684–717.

Roy, Raja Devasish, "Country technical notes on indigenous peoples' issues: People's Republic of Bangladesh," Technical Report 2012.

Scacco, Alexandra and Shana S Warren, "Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria," *American Political Science Review*, 2018, *112* (3), 654–677.

Schaller, Mark and Robert B Cialdini, "The economics of empathic helping: Support for a mood management motive," *Journal of Experimental Social Psychology*, 1988, *24* (2), 163–181.

Schwartz, Shalom H and Wolfgang Bilsky, "Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications.," *Journal of Personality and Social Psychology*, 1990, *58* (5), 878.

Schwarz, Norbert and Herbert Bless, "Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning," in "Emotion and Social Judgements," Garland Science, 1991, pp. 55–71.

Selten, Reinhard and Axel Ockenfels, "An experimental solidarity game," *Journal of Economic Behavior & Organization*, 1998, *34* (4), 517–539.

Shamsuddoha, Md and Ms Rifat Jahan, "Santal community in Bangladesh: A socio-historical analysis," *Asian Journal of Humanity, Art and Literature*, 2018, *5* (2), 89–100.

Siddique, Abu and Michael Vlassopoulos, "Competitive preferences and ethnicity: Experimental evidence from Bangladesh," *The Economic Journal*, 2020, *130* (627), 793–821.

_ , _ , and Yves Zenou, "Market Competition and Discrimination," *European Economic Review*, 2023, *152*, 104361.

_ , _ , and _ , "Replication Data for: 'Leveraging Virtual Contact and Social Networks to Foster Interethnic Harmony'," 2025. Harvard Dataverse, https://doi.org/10.7910/DVN/WJAOOA.

Siraj, Nasrin and Ellen Bal, "'Hunger has brought us into this jungle': Uderstanding mobility and immobility of Bengali immigrants in the Chittagong Hills of Bangladesh," in "Patterns of Im/mobility, Conflict and Identity," Routledge, 2021, pp. 18–34.

Sze, Jocelyn A, Anett Gyurak, Madeleine S Goodkind, and Robert W Levenson, "Greater emotional empathy and prosocial behavior in late life.," *Emotion*, 2012, *12* (5), 1129.

Tanguy, Bernard, Stefan Dercon, Kate Orkin, and Alemayehu Seyoum Taffesse, "The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia," Technical Report 2014.

The Business Standard, "Santal farmers committed suicide after irrigation water was denied: Police," https://www.tbsnews.net/bangladesh/santal-farmers-committed-suicide-after-irrigation-water-was-denied-police-418422 2022. Accessed: 2025-01-30.

The Daily Star, "Discrimination and marginalisation of ethnic minorities," https://www.thedailystar.net/editorial/discrimination-and-marginalisation-ethnic-minorities-1267645 2016. Accessed: 2025-01-30.

_ , "Plainland Ethnic Minority Communities: Nearly 80% of households have no access to credit: report," https://www.thedailystar.net/business/news/nearly-80-households-have-no-access-credit-report-3605261 2024. Accessed: 2025-01-30.

Tropp, Linda R and Thomas F Pettigrew, "Differential relationships between intergroup contact and affective and cognitive dimensions of prejudice," *Personality and Social Psychology Bulletin*, 2005, *31* (8), 1145–1158.

Turner, Rhiannon N, Richard J Crisp, and Emily Lambert, "Imagining intergroup contact can improve intergroup attitudes," *Group Processes & Intergroup Relations*, 2007, *10* (4), 427–441.

Uddin, Md Emaj, "Cross-cultural socio-economic status attainment between Muslim and Santal couple in rural Bangladesh," *International Journal of Humanities and Social Sciences*, 2009, *3* (7), 1512–1519.

Ushchev, Philip and Yves Zenou, "Social norms in networks," *Journal of Economic Theory*, 2020, *185*, 104969.

Valente, Thomas W, "Network interventions," *Science*, 2012, *337* (6090), 49–53.

Veselovsky, Veniamin, Manoel Horta Ribeiro, and Robert West, "Artificial intelligence: Crowd workers widely use large language models for text production tasks," 2023. arXiv preprint arXiv:2306.07899.

Wang, Tianyi, "Media, pulpit, and populist persuasion: Evidence from Father Coughlin," *American Economic Review*, 2021, *111* (9), 3064–3092.

Watson, Goodwin, "Resistance to change," *American Behavioral Scientist*, 1971, *14* (5), 745–766.

Weiss, Chagai M, Shira Ran, and Eran Halperin, "Educating for inclusion: Diversity education programs can reduce prejudice toward outgroups in Israel," *Proceedings of the National Academy of Sciences*, 2023, *120* (16), e2218621120.

Westfall, Peter H and S Stanley Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Vol. 279, John Wiley & Sons, 1993.

Yanagizawa-Drott, David, "Propaganda and conflict: Evidence from the Rwandan genocide," *Quarterly Journal of Economics*, 2014, *129* (4), 1947–1994.

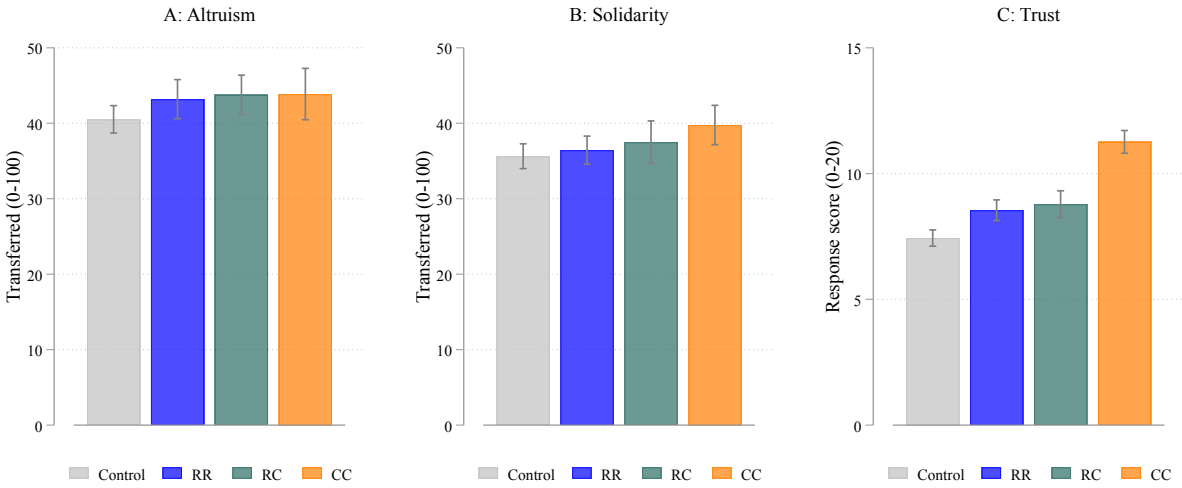
Young, Alwyn, "Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results," *Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Zárate, Román Andrés, "Uncovering peer effects in social and academic skills," *American Economic Journal: Applied Economics*, 2023, *15* (3), 35–79.

Zeelenberg, Marcel, Rob MA Nelissen, Seger M Breugelmans, and Rik Pieters, "On emotion specificity in decision making: Why feeling is for doing," *Judgment and Decision Making*, 2008, *3* (1), 18–27.

Figures

Figure I: Prosocial outcomes at endline



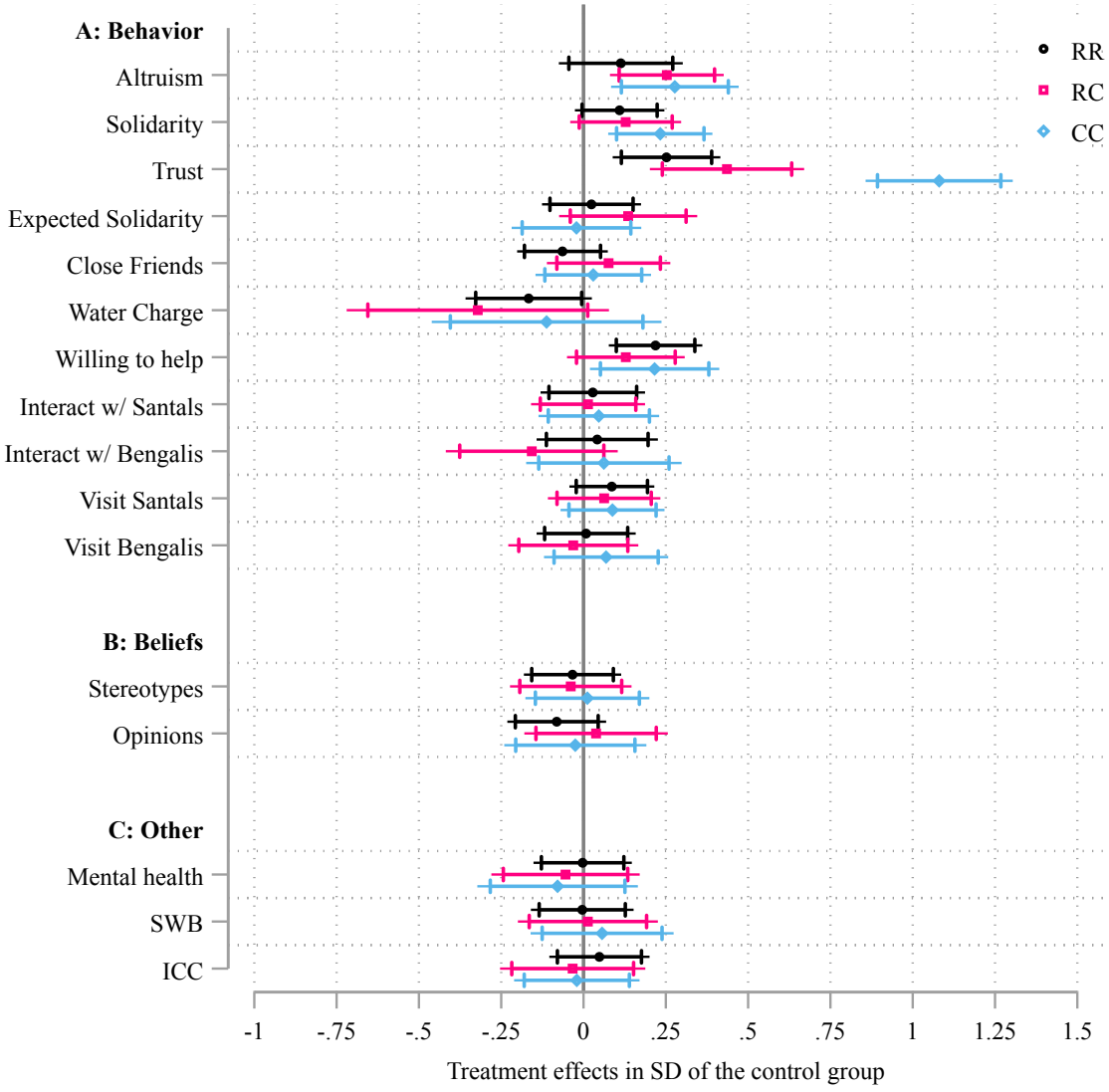
Note: The sample includes treated Bengalis in the two treatment arms and the control arm. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level. All p-values below are unadjusted two-sided tests on raw means.

Graph A: Control vs RR: $p = 0.10$; Control vs RC: $p = 0.05$; Control vs CC: $p = 0.09$; RR vs RC: $p = 0.76$; RR vs CC: $p = 0.76$; RC vs CC: $p = 0.96$.

Graph B: Control vs RR: $p = 0.52$; Control vs RC: $p = 0.26$; Control vs CC: $p = 0.01$; RR vs RC: $p = 0.54$; RR vs CC: $p = 0.04$; RC vs CC: $p = 0.22$.

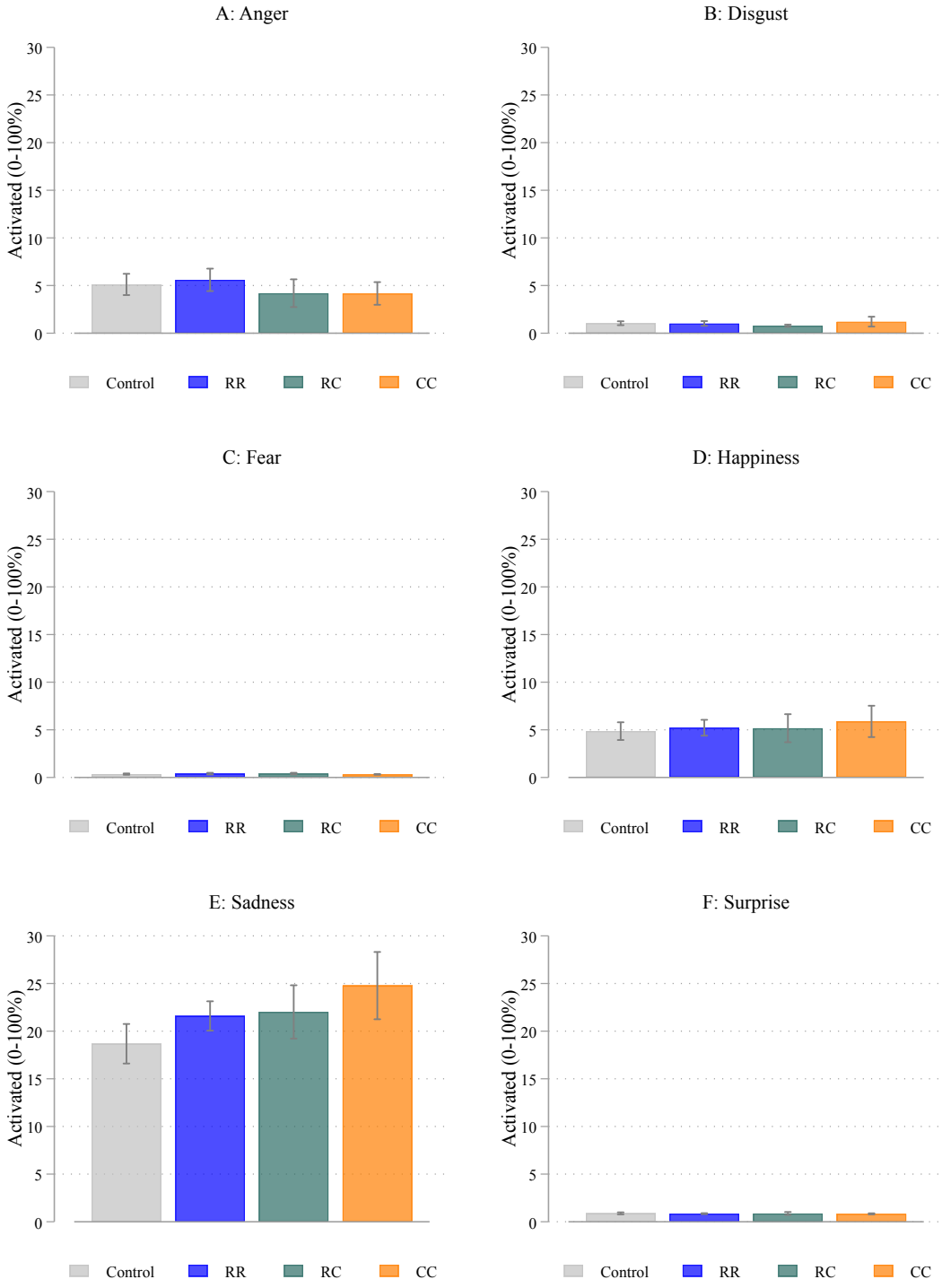
Graph C: Control vs RR: $p < 0.0001$; Control vs RC: $p < 0.0001$; Control vs CC: $p < 0.0001$; RR vs RC: $p = 0.49$; RR vs CC: $p < 0.0001$; RC vs CC: $p < 0.0001$.

Figure II: Treatment effects on behavior, beliefs, and other outcomes of the treated ethnic majority sample



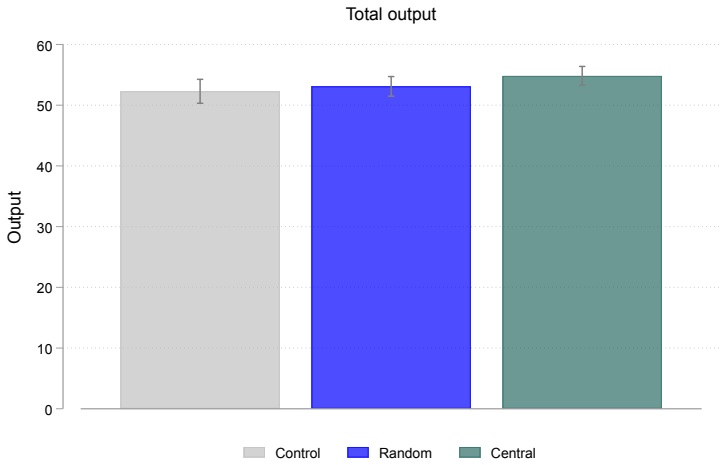
Note: This figure reports treatment effects in SD units of the control group, with 90% and 95% confidence intervals. All outcomes have been control-group standardized.

Figure III: Raw differences in emotions of treated ethnic majority

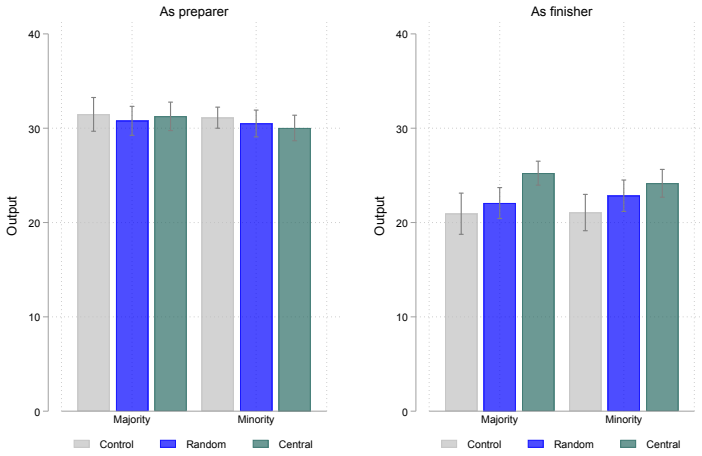


Note: The sample includes treated Bengalis in both treatment arms and the control arm. The sum of all six emotions reported above, along with the neutral emotion (i.e., no emotions), equals 100. This means that each individual's emotions were scored between 0-100. A higher value indicates a stronger presence of that particular emotion. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level. ***Graph E:*** Control vs RR: $p = 0.03$; Control vs RC: $p = 0.06$; Control vs CC: $p < 0.01$; RR vs RC: $p = 0.80$; RR vs CC: $p = 0.11$; RC vs CC: $p = 0.21$.

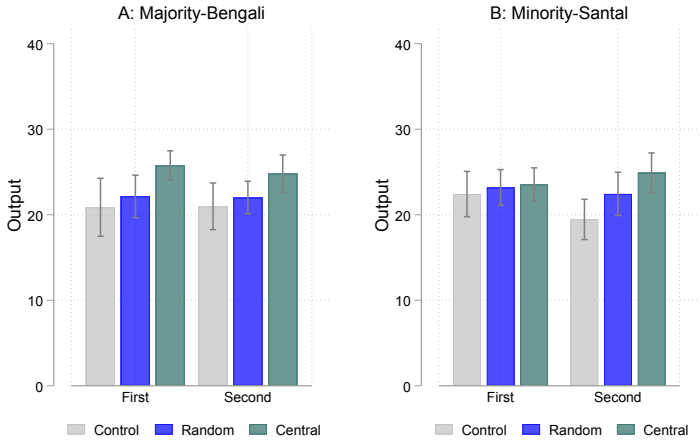
Figure IV: Casual work field experiment



Panel A: Total casual work output, by treatment



Panel B: Total output, by ethnicity and treatment



Panel C: Output as finisher, by ethnicity

Note: Bars show mean output. Error bars report 95% confidence intervals with standard errors clustered at the village level. Comparisons are unconditional (no covariate adjustment). *Panel A:* Total output is preparer+finisher. Control vs Random: $p = 0.53$; Control vs Central: $p = 0.05$; Random vs Central: $p = 0.13$. *Panel B:* Majority=Bengali, Minority=Santal. Majority preparer: $p = 0.57, 0.86, 0.67$; Minority preparer: $p = 0.50, 0.22, 0.63$; Majority finisher: $p = 0.42, < 0.01, < 0.01$; Minority finisher: $p = 0.17, 0.01, 0.25$ (Control vs Random, Control vs Central, Random vs Central). *Panel C:* 'First' denotes assigned finisher at start; 'Second' denotes completing as finisher. Majority (First): $p = 0.56, 0.01, 0.02$; Majority (Second): $p = 0.55, 0.04, 0.06$; Minority (First): $p = 0.65, 0.50, 0.81$; Minority (Second): $p = 0.09, < 0.01, 0.16$ (Control vs Random, Control vs Central, Random vs Central).

Tables

Table I: Treatment effects on behaviors of the treated ethnic majority

	Prosociality			Self-reports				Interactions & visits			
	Altruism	Solidarity	Trust	Ex Sol	Friends	Water$	Help	Int San	Int Ben	Visit San	Visit Ben
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
RR	2.593	2.525*	0.878***	0.451	-0.096	-0.849*	0.420***	0.023	0.007	1.231	0.276
	(2.218)	(1.475)	(0.318)	(1.694)	(0.117)	(0.498)	(0.147)	(0.074)	(0.015)	(0.834)	(1.340)
RC	5.966***	2.867	1.617***	3.027	0.138	-1.633	0.253	0.012	-0.025	0.869	-0.469
	(2.076)	(1.869)	(0.454)	(2.403)	(0.162)	(1.031)	(0.186)	(0.082)	(0.022)	(1.110)	(1.781)
CC	6.542***	5.178***	4.063***	-0.527	0.059	-0.566	0.432**	0.044	0.011	1.192	1.306
	(2.336)	(1.755)	(0.433)	(2.253)	(0.152)	(0.907)	(0.204)	(0.086)	(0.020)	(1.026)	(1.702)
Observations	1,516	1,515	1,513	1,511	1,519	444	1,394	1,515	1,515	1,515	1,514
Control mean	40.52	35.63	7.437	29.84	1.861	0.521	6.575	1.789	3.985	19.30	48.58
Control SD	23.77	21.95	3.796	22.62	1.712	5.103	2.045	0.932	0.163	12.80	17.80
RR=RC p-values	0.13	0.85	0.09	0.29	0.14	0.28	0.30	0.89	0.14	0.72	0.67
RR=CC p-values	0.12	0.11	<0.001	0.67	0.29	0.69	0.95	0.81	0.87	0.97	0.54
RC=CC p-values	0.77	0.19	<0.001	0.08	0.58	0.26	0.24	0.73	0.15	0.77	0.25
RR+RC vs C p-values	0.08	0.06	<0.01	0.46	0.80	0.09	<0.01	0.77	0.90	0.16	0.96
FWER *p*-values on RR	0.11	0.04	<0.01	0.97	0.90	0.97	0.05	0.99	0.95	0.65	0.99
FWER *p*-values on RC	<0.01	0.11	<0.001	0.64	0.84	0.88	0.64	0.99	0.51	0.89	0.99
FWER *p*-values on CC	<0.001	<0.01	<0.001	0.97	0.97	0.97	0.27	0.91	0.99	0.60	0.99
RI *p*-values on RR	0.26	0.07	<0.01	0.69	0.45	0.82	<0.01	0.72	0.50	0.21	0.92
RI *p*-values on RC	0.01	0.20	<0.01	0.20	0.30	0.44	0.20	0.74	0.16	0.35	0.65
RI *p*-values on CC	<0.01	<0.01	<0.001	0.98	0.60	0.83	0.08	0.37	0.87	0.18	0.68

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in both treatment arms and the control arm. RR is the treatment arm where participants were selected randomly and are from the *Random* arm; RC is the sample where participants were selected randomly and are from the *Central* arm; CC is the sample where all participants had high diffusion centrality, were selected using the 'gossip' method, and are from the *Central* arm. All specifications include union council fixed effects, outcome measured at baseline, education, income, and other controls selected using the post-double selected LASSO (Belloni et al., 2014). Outcomes in the columns are as follows: (1) Altruism or the amount (0-100) transferred to a minority measured using the dictator game (Eckel and Grossman, 1996); (2) Solidarity or the amount (0-100) transferred to a minority, conditional on a risk shock, measured using a simplified solidarity game (Selten and Ockenfels, 1998); (3) Trust is the level of trust towards minorities measured using a slightly modified Global Preference Survey question (Falk et al., 2018), which was answered on a 0-20 thermometer scale; (4) Ex Sol is the solidarity expected from the opponent (0-100); (5) Friends is the number of minority friends among the ten closest friends; (6) Water$ is the amount charged to minorities in BDT when they come to fetch water from own tubewell; (7) Help is the participant's willingness to help others (answered on a scale between 0-10, where 10 means very willing to help); (8-9) social interactions with non-coethnics and coethnics; (10-11) frequency of visits to non-coethnics' and coethnics' house. FWER *p*-values correspond to multiple hypotheses testing-adjusted *p*-values computed using the Westfall and Young (1993) corrections. To compute the FWER-adjusted *p*-values, we consider prosociality, self-reported behaviors, and interactions & visits as three families of outcomes. RI *p*-values correspond to randomization inference based *p-values* computed following Young (2019). We used OLS regressions (with all controls available), instead of 'pdslasso' to compute the FWER and RI *p*-values. We also report the control group's means and standard deviations. 'RR+RC vs C p-values' report the p-values of the pooled effects of RR and RC (where C is the 'Control' group).

Table II: Impacts on beliefs of treated ethnic majority

	Panel A: Stereotypes						
	Stereotype Index	**Unclean**	**Doctors**	**Establish**	**Teachers**	**Education**	**Agriculture**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
RR	-0.033	0.216	0.238***	0.150	-0.077	-0.100	-0.628**
	(0.075)	(0.148)	(0.063)	(0.111)	(0.251)	(0.176)	(0.297)
RC	-0.039	0.005	-0.026	-0.004	0.090	-0.224	-0.127
	(0.094)	(0.187)	(0.089)	(0.119)	(0.363)	(0.227)	(0.371)
CC	0.012	0.150	0.012	-0.017	0.121	-0.024	-0.035
	(0.096)	(0.183)	(0.084)	(0.130)	(0.345)	(0.234)	(0.372)
Observations	1,476	1,503	1,500	1,496	1,502	1,498	1,495
Control mean	0	2.738	3.241	3.875	4.820	3.230	5.530
Control SD	1	2.024	1.016	1.376	3.459	2.584	3.993
RR=RC p-values	0.95	0.24	<0.01	0.18	0.63	0.57	0.16
RR=CC p-values	0.62	0.71	<0.01	0.19	0.55	0.74	0.10
RC=CC p-values	0.56	0.41	0.61	0.92	0.92	0.36	0.79
RR+RC vs C p-values	0.62	0.27	<0.01	0.29	0.90	0.42	0.07
FWER p-values on RR	1.00	0.56	<0.001	0.58	1.00	1.00	0.20
FWER p-values on RC	1.00	1.00	1.00	1.00	1.00	0.93	1.00
FWER p-values on CC	1.00	1.00	1.00	1.00	1.00	1.00	1.00
RI p-values on RR	0.71	0.13	<0.001	0.14	0.79	0.64	0.04
RI p-values on RC	0.58	0.89	0.73	0.57	0.84	0.32	0.74
RI p-values on CC	0.99	0.48	0.97	0.67	0.71	0.85	0.95

	Panel B: Discriminatory Opinions							
	Discriminatory Index	**Honest**	**Eat**	**Schools**	**Teaching**	**Working**	**Child's fr**	**Trust**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
RR	-0.081	0.020	-0.057	-0.064	0.001	-0.165	-0.120	0.143*
	(0.076)	(0.063)	(0.153)	(0.237)	(0.108)	(0.234)	(0.107)	(0.078)
RC	0.038	0.013	-0.004	-0.035	0.048	0.218	-0.096	-0.160
	(0.111)	(0.077)	(0.170)	(0.327)	(0.154)	(0.323)	(0.135)	(0.128)
CC	-0.025	0.171**	-0.175	-0.002	-0.082	0.095	-0.206	-0.051
	(0.110)	(0.078)	(0.185)	(0.320)	(0.165)	(0.332)	(0.138)	(0.121)
Observations	1,485	1,513	1,511	1,511	1,505	1,498	1,500	1,501
Control mean	0	4.282	7.268	5.019	5.096	5.347	5.364	3.586
Control SD	1	0.895	1.927	3.162	1.634	2.906	1.427	1.178
RR=RC p-values	0.25	0.94	0.73	0.93	0.72	0.21	0.86	0.01
RR=CC p-values	0.58	0.07	0.49	0.84	0.57	0.41	0.54	0.09
RC=CC p-values	0.51	0.02	0.35	0.91	0.42	0.66	0.39	0.30
RR+RC vs C p-values	0.51	0.75	0.77	0.80	0.90	0.79	0.24	0.45
FWER p-values on RR	0.87	1.00	1.00	1.00	1.00	1.00	0.72	0.39
FWER p-values on RC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.91
FWER p-values on CC	1.00	0.30	0.97	1.00	1.00	1.00	0.79	1.00
RI p-values on RR	0.22	0.87	0.74	0.77	0.98	0.47	0.14	0.07
RI p-values on RC	0.98	0.91	0.98	0.84	0.91	0.67	0.52	0.27
RI p-values on CC	0.66	0.06	0.36	0.99	0.55	0.78	0.20	0.71

Robust Standard Errors clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in both treatment arms and the control arm. See the note under Table I for specification details and p-value definitions. All other information is as previously described. In this table, only the 'Index' outcomes (column 1) have been control group standardized, where the mean of the control group is 0 and SD is 1. Outcomes in the other columns are as follows: **Panel A**: (2) Santals are often unclean; (3) Santals would not make good doctors; (4) I do not know any Santals who have established themselves; (5) Santals do not make very good school teachers; (6) Santals do not continue beyond schools; (7) Santals should continue working in the agriculture sector. **Panel B:** (2) All Santals I know are not honest people; (3) I never eat food and drinks offered by Santals; (4) schools should be separate for Santals and Bengalis; (5) there should not be more Santal teachers in my child's school; (6) I do not enjoy working/doing business with Santals; (7) most of my child's best friends are Bengalis; (8) one can easily trust a Santal (for presentation purposes, we report this here in non-reversed).

Table III: Treatment effects on outcomes of the untreated ethnic majority

Panel A: Behaviors

	Prosociality			Self-reports			Interactions & visits			
	Altruism	**Solidarity**	**Trust**	**Ex Sol**	**Friends**	**Water$**	**Int San**	**Int Ben**	**Visit San**	**Visit Ben**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Random (R)	1.245	-2.118	0.289	-3.370	0.212	-0.474	0.043	-0.013	1.187*	1.386
	(3.331)	(4.699)	(0.397)	(3.446)	(0.221)	(0.333)	(0.112)	(0.055)	(0.673)	(1.373)
Central (C)	6.979*	9.958**	0.679**	-8.239	1.165***	-1.045**	-0.140	0.124**	0.894	2.790*
	(3.642)	(4.168)	(0.342)	(6.108)	(0.298)	(0.454)	(0.130)	(0.061)	(1.029)	(1.547)
Observations	800	800	797	794	800	214	800	799	799	800
Control mean	33.45	36.50	8.347	40.77	2.625	0.909	2.094	3.585	5.094	10.93
Control SD	21.80	22.49	3.980	21.40	1.252	3.057	0.588	0.674	2.014	3.441
R=C p-values	0.10	<0.01	0.27	0.42	<0.001	0.23	0.18	0.01	0.79	0.35
FWER *p*-values on R	0.89	0.89	0.89	0.36	0.39	0.39	0.81	0.91	0.06	0.46
FWER *p*-values on C	0.15	0.01	0.15	0.30	0.04	0.30	0.46	0.03	0.71	0.11

Panel B: Beliefs

	Stereotypes	**Opinions**
Variables	**(1)**	**(2)**
Random (R)	0.004	-0.293**
	(0.099)	(0.146)
Central (C)	-0.031	0.526***
	(0.123)	(0.192)
Observations	780	784
R=C *p*-values	0.78	<0.01
FWER *p*-values on R	0.95	0.15
FWER *p*-values on C	0.95	<0.01

Note: Robust standard errors clustered at village level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The sample includes untreated Bengalis in the two treatment arms and in control villages. Random (R) denotes untreated Bengalis in *Random* villages and Central (C) denotes untreated Bengalis in *Central* villages. See the note under Table I for specification details and outcome definitions. For FWER adjustment, we treat prosociality, self-reported behaviors, and interactions & visits as three outcome families (Panel A), and stereotypes and discriminatory opinions as two outcome families (Panel B). Panel B outcomes are standardized relative to the control group (mean 0, SD 1); for both indices, negative coefficients indicate more favorable views of non-coethnics.

Table IV: Treatment effects on behaviors of the untreated ethnic minority

	Prosociality			Self-reports			Interactions & visits			
	Altruism	**Solidarity**	**Trust**	**Ex Sol**	**Friends**	**Water$**	**Int San**	**Int Ben**	**Visit San**	**Visit Ben**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Random (R)	0.895	4.846	0.765**	-3.695	0.358	0.507	-1.036**	0.094	23.712***	18.717***
	(4.508)	(4.198)	(0.343)	(3.047)	(0.392)	(6.038)	(0.522)	(0.396)	(4.272)	(4.434)
Central (C)	3.618	6.112	1.504***	4.365	-0.386	-1.859	-0.288	0.588	22.900***	10.584**
	(4.583)	(5.000)	(0.359)	(3.747)	(0.428)	(6.732)	(0.577)	(0.422)	(5.192)	(5.309)
Observations	721	721	717	720	721	93	720	719	720	721
Control mean	38.63	35.52	6.930	40.72	3.532	18.16	2.766	1.463	24.12	18.63
Control SD	21.99	23.01	3.818	22.46	1.732	10.70	1.817	0.995	18.77	19.81
R=C p-value	0.52	0.72	<0.01	0.01	0.03	0.82	0.12	0.25	0.82	<0.01
FWER p-values on R	0.57	0.21	<0.01	0.65	0.77	0.85	<0.01	0.81	<0.001	<0.001
FWER p-values on C	0.21	0.12	<0.001	0.44	0.82	0.85	0.48	0.02	<0.001	<0.01

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes untreated Santals in both treatment arms and the control arm. Random (R) denotes untreated Santals in Random villages and Central (C) denotes untreated Santals in Central villages. See the note under Table I for specification details, and p-value and outcome definitions. To compute the FWER-adjusted p-values, we consider prosociality, self-reported behaviors, and interactions & visits as three families of outcomes. All other information is as previously described.

Table V: Treatment effects on interethnic complaints using administrative data

	To village counselors		To police stations	
	by Santal	by Bengali	by Santal	by Bengali
Variables	**(1)**	**(2)**	**(3)**	**(4)**
Random (R)	0.117	0.156	0.026	-0.130
	(0.278)	(0.480)	(0.230)	(0.277)
Central (C)	0.102	0.063	0.082	-0.655**
	(0.334)	(0.562)	(0.337)	(0.326)
Control mean	1.05	1.72	0.90	1.58
	[0.86]	[1.54]	[0.84]	[0.93]
Controls:				
Village size	Y	Y	Y	Y
Village ethnic diversity	Y	Y	Y	Y
Complaints at baseline	Y	Y	Y	Y
Union Council FE	Y	Y	Y	Y
R=C p-values	0.96	0.85	0.86	0.16
Observations	115	115	117	117
R-squared	0.154	0.229	0.150	0.247
FWER p-values on R	1.00	1.00	1.00	1.00
FWER p-values on C	1.00	1.00	1.00	0.35

Robust standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table reports treatment effects on interethnic dispute complaints in a month at the village level, i.e., Santals filing complaints against Bengalis and vice versa. Therefore, the outcomes in all columns are the number of complaints filed against non-coethnics. Columns 1 and 3 report complaints made by ethnic minority Santals against ethnic majority Bengali; and Columns 2 and 4 report complaints made by ethnic majority Bengalis against ethnic minority Santals. Complaint records from three villages were missing at the village counselors' offices, and other records at those offices or police stations were unconfirmed due to large discrepancies between village names used by residents and official records. Village size is the total number of households per village, and village ethnic diversity is the proportion of Santal households per village. FWER p-values correspond to multiple hypotheses testing-adjusted p-values computed using the Westfall and Young (1993) corrections. To compute the FWER-adjusted p-values, we consider this a single family of outcomes.

Leveraging Virtual Contact and Social Networks to Foster Interethnic Harmony

Online Appendix

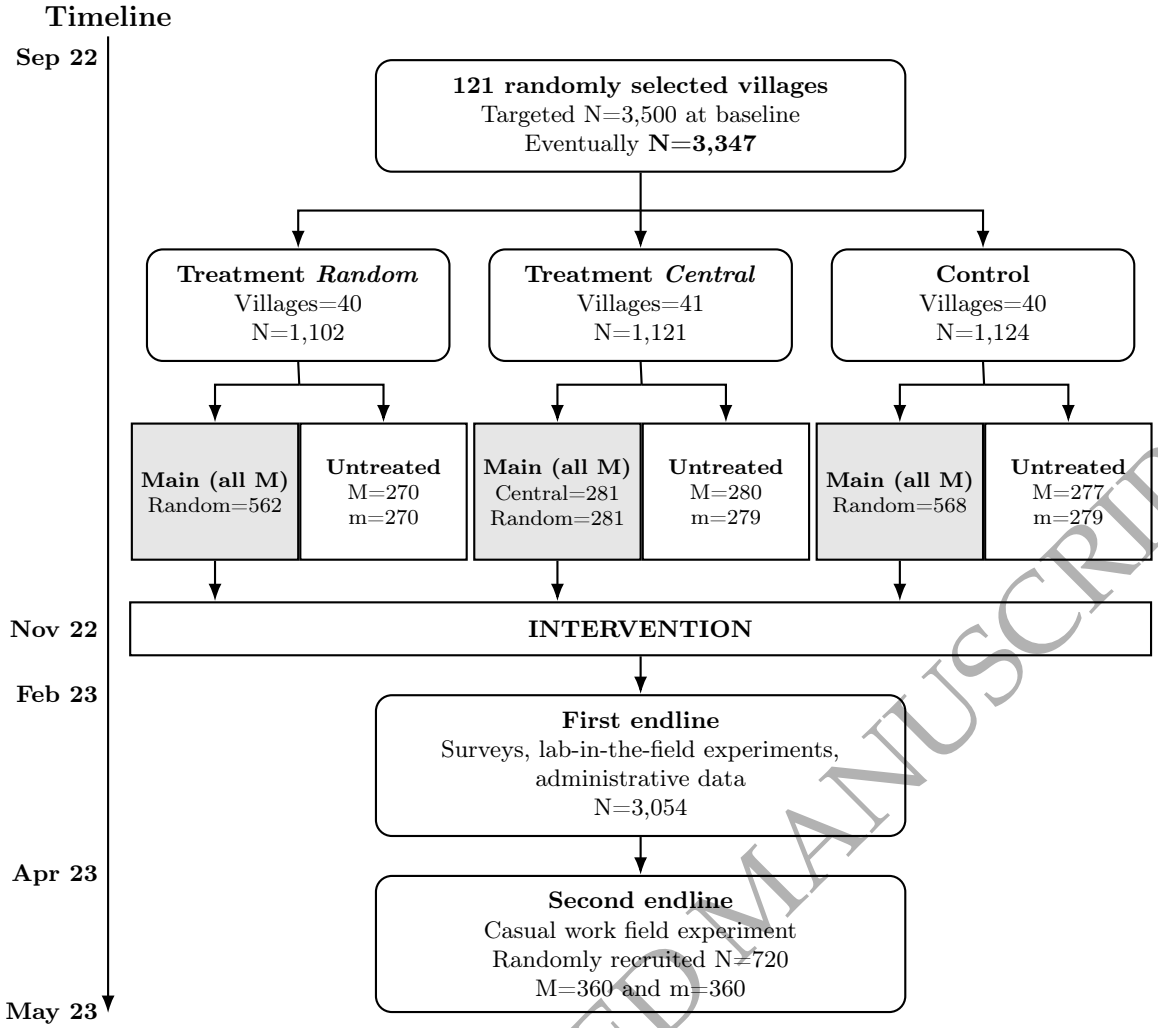
By Abu Siddique, Michael Vlassopoulos, Yves Zenou

A Additional Figures and Tables

Figure A1: The documentary film

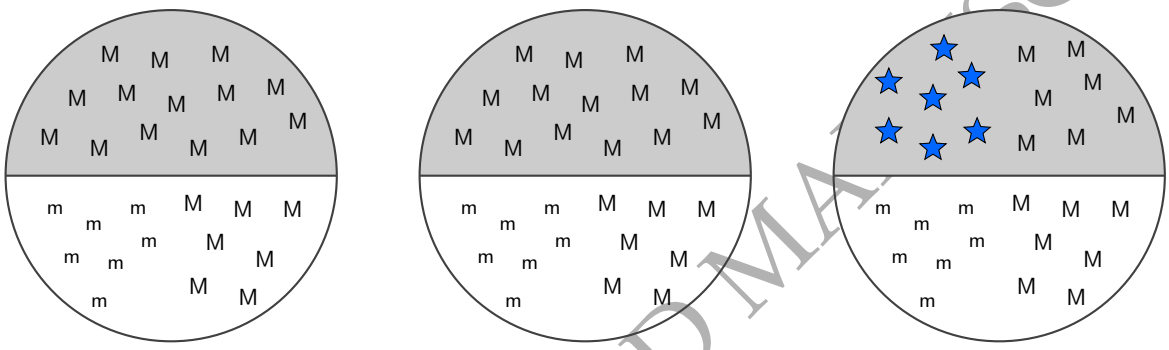
Note: The full film can be accessed here: `https://www.youtube.com/watch?v=hWizDrLXLoc`.

Figure A2: Study flowchart



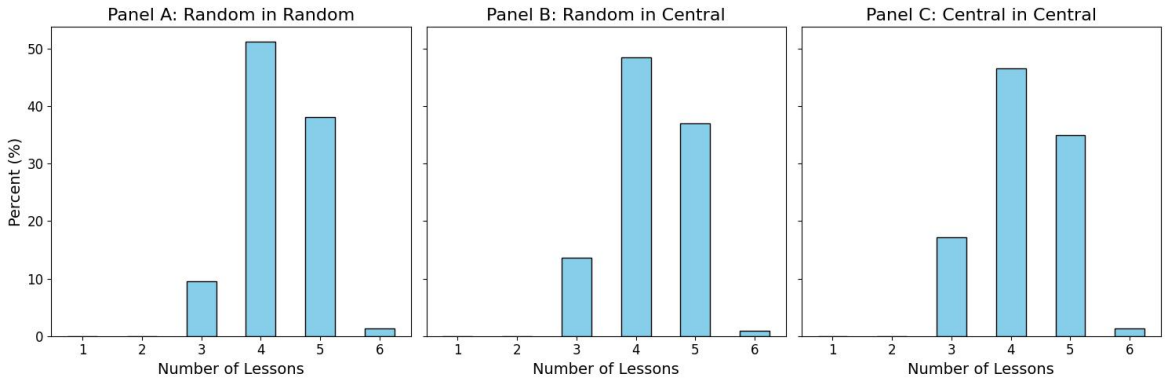
Note: *M* denotes the ethnic majority Bengalis and *m* the ethnic minority Santals. *Random* denotes the participants that were randomly selected, whereas *Central* denotes those selected following the approach in Banerjee et al. (2019a). *N* is the sample size. Centrality was measured only in Central villages; CC were excluded from the random draw; untreated rosters were set before screenings.

Figure A3: Sampling



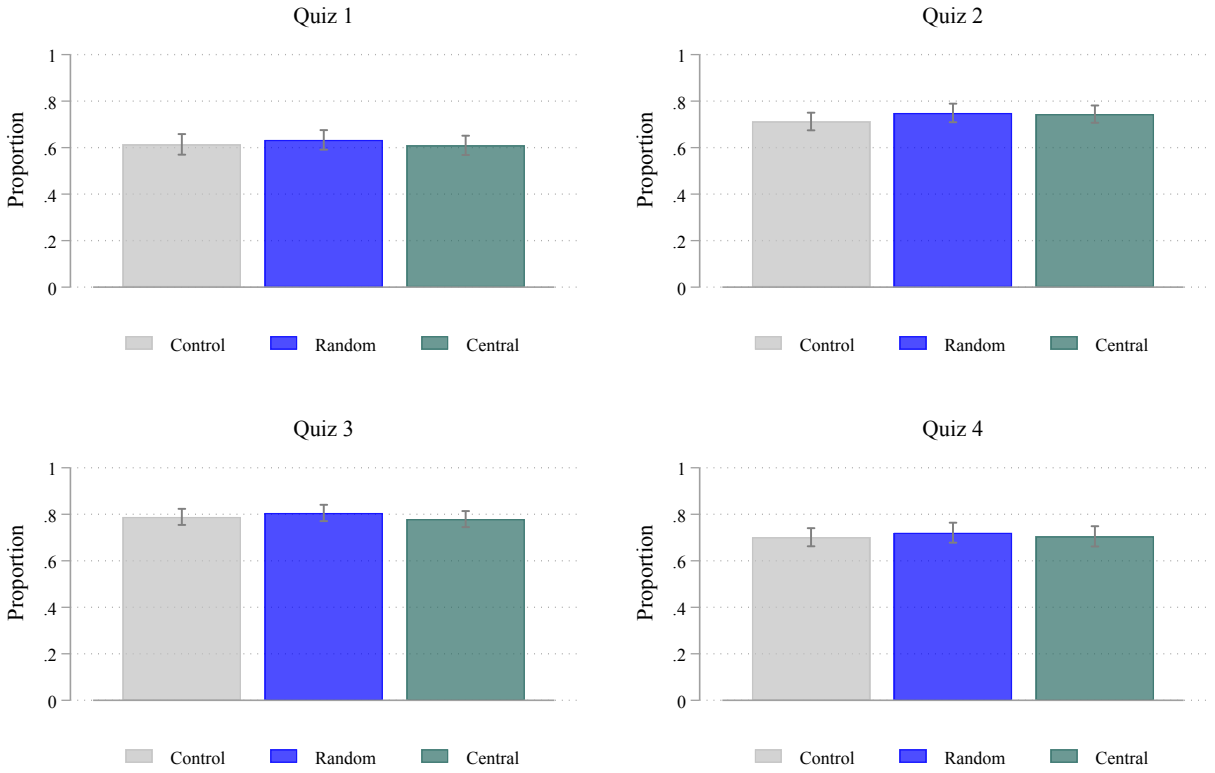
Note: M = ethnic majority, m = ethnic minority, $\star$ = ethnic majority with high diffusion centrality. Diffusion centrality is measured only in Central villages for the seven pre-identified majority households per village. The shaded region includes treated participants who watched the documentary film, while the white region includes untreated participants.

Figure A4: Frequency of new information acquisition among treated ethnic majority



Note: The figure displays the distribution of the number of new information participants reported learning from the documentary, in response to the question: "Name five new things you learned from the video". The sample includes treated Bengali participants across two main treatment arms. Panel A shows the results for the 'Random in Random' or simply *Random* treatment arm. Panels B ('Random in Central') and C ('Central in Central') show results for two subgroups that together constitute the second (*Central*) treatment arm. The percentages on the y-axis are calculated within each panel's respective sample. For example, the bars in Panel B represent the percentage of respondents only from the 'Random in Central' subgroup, not from the entire *Central* treatment arm. Note that the sample size in the 'Random in Random' arm (Panel A) is approximately equal to the combined sample sizes of the 'Random in Central' and 'Central in Central' subgroups (Panels B and C). Note also that some respondents mentioned a sixth new piece of information learned, and enumerators recorded that response, which is why there are six bars.

Figure A5: Answers to post-film quiz

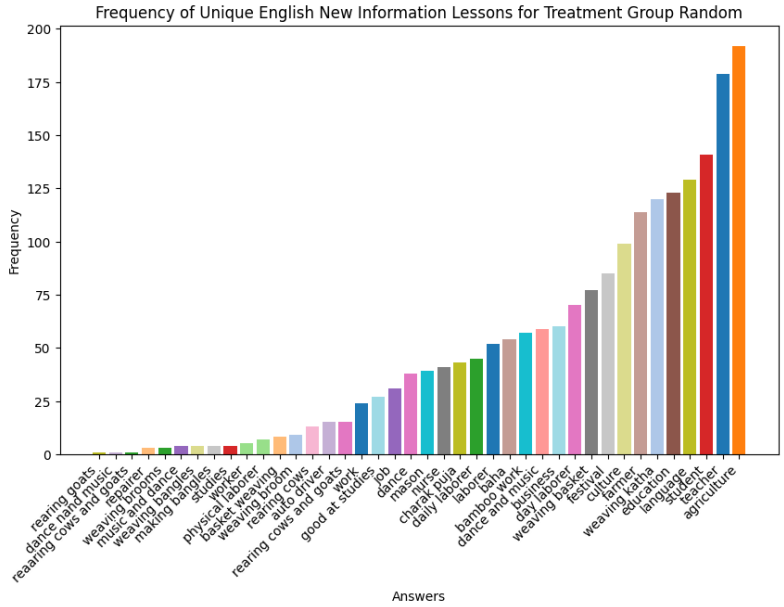


Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. Each bar shows the proportion of respondents who got the quiz correct. Quiz questions were open-ended, and each response was recorded as either "correct" or "incorrect". Bars indicate 95% confidence intervals, with standard errors clustered at the village level. The quiz questions for the treatment group were as follows: (Quiz 1) In the film, you watched a festival celebrated by the Santals. What is the name of that festival (*Answer: Baha festival*)? (Quiz 2) What are the main reasons households frequently require repairs (*Answer: heavy rain/storm/cyclone*)? (Quiz 3) You saw a Santal college teacher in the film. What is the name of the college where this teacher works (*Answer: Rajshahi college*)? (Quiz 4) Some secondary school girls shared their future aspirations in the film. Can you tell me what they mostly aspire to become (*Answer: Doctor/nurse*)? Quiz questions for the control group were as follows: (Quiz 1) In the film, you watched a flower farming nursery. What is the name of that flower farming business (*Answer: Mother's Blessing*)? (Quiz 2) What are the main reasons flower farming frequently experience loss (*Answer: heavy rain/storm/cyclone*)? (Quiz 3) You saw the farmer talking about his four daughters in the film. What is the name of the college where his oldest daughter studies (*Answer: Rajshahi college*)? (Quiz 4) According to the farmer, when most of the flowers get sold (*Answer: Bengali New Year/Valentine's Day*)?

Figure A6: Measuring emotions

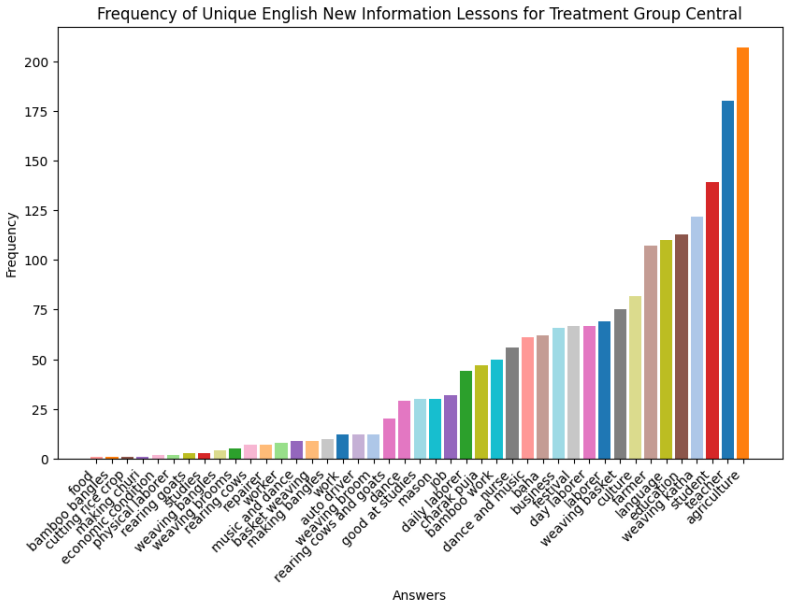
Note: Measuring emotions using pictures.

Figure A7: Histogram of most frequent things treated ethnic majority learned in Random arm only



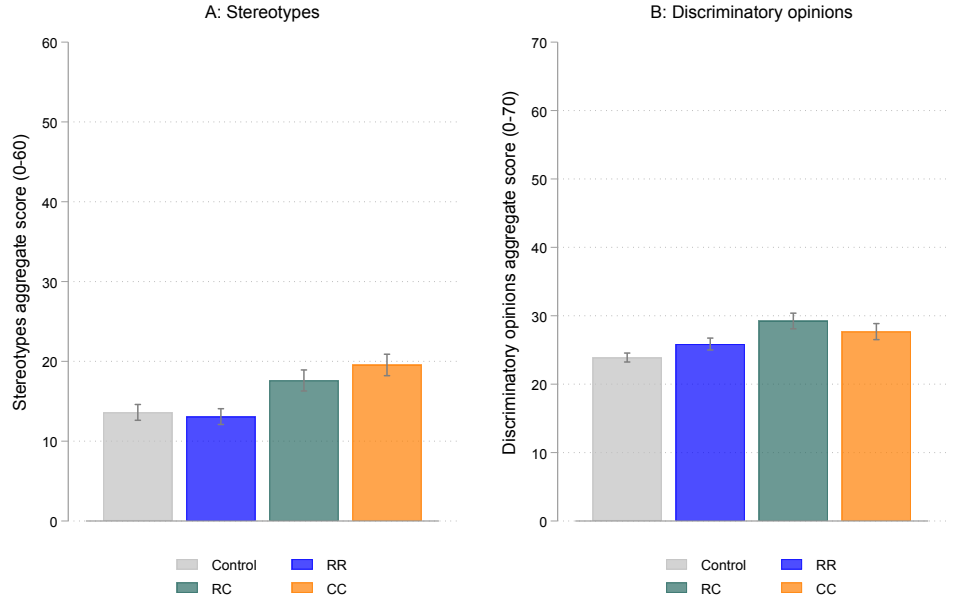
Note: The sample includes treated Bengalis in Random arm.

Figure A8: Histogram of most frequent things treated ethnic majority learned in Central arm only



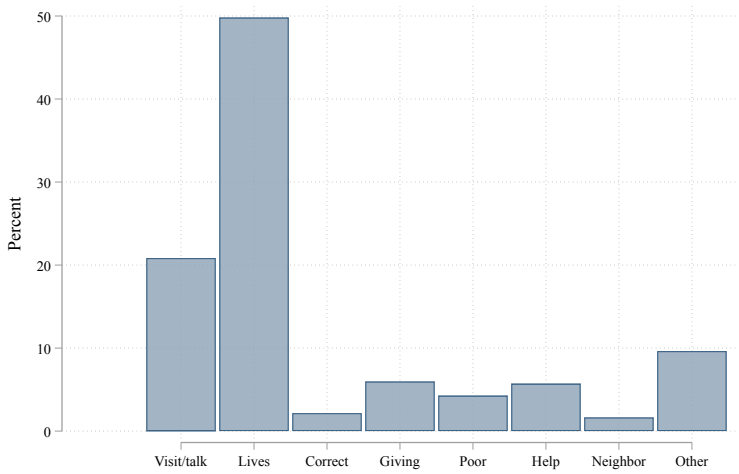
Note: The sample includes treated Bengalis in Central arm.

Figure A9: Differences in stereotypes and discriminatory opinions at baseline



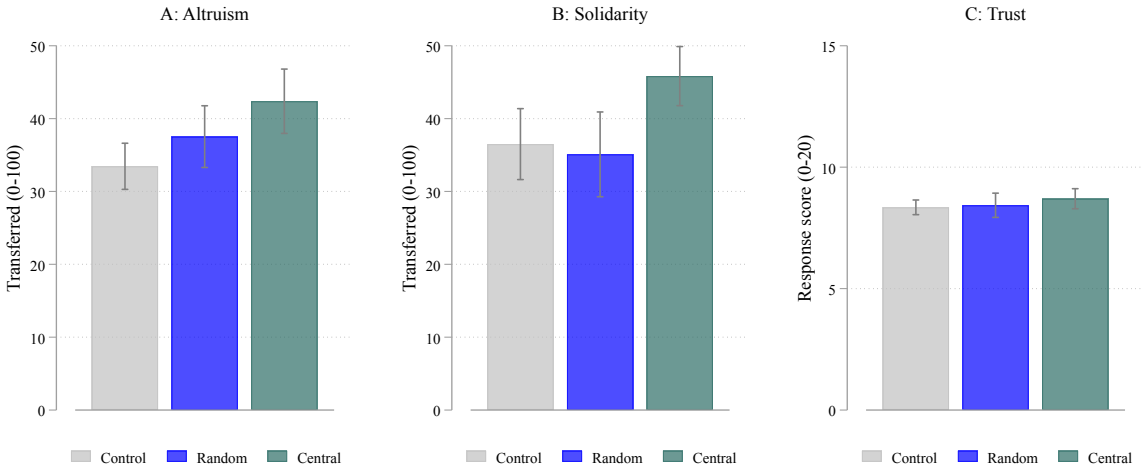
Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. Stereotypes is a score between 0-60, and Discriminatory opinions is a score between 0-70 (both explained in detail in Appendix D.1), where a higher number corresponds to having more negative stereotypes or discriminatory opinions. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level:
Graph A: Control vs RR: $p = 0.80$; Control vs RC: $p = 0.08$; Control vs CC: $p < 0.01$; RR vs RC: $p = 0.02$; RR vs CC: $p < 0.01$; RC vs CC: $p = 0.31$.
Graph B: Control vs RR: $p = 0.07$; Control vs RC: $p < 0.001$; Control vs CC: $p < 0.01$; RR vs RC: $p < 0.01$; RR vs CC: $p = 0.10$; RC vs CC: $p = 0.21$.

Figure A10: Perceived study purpose



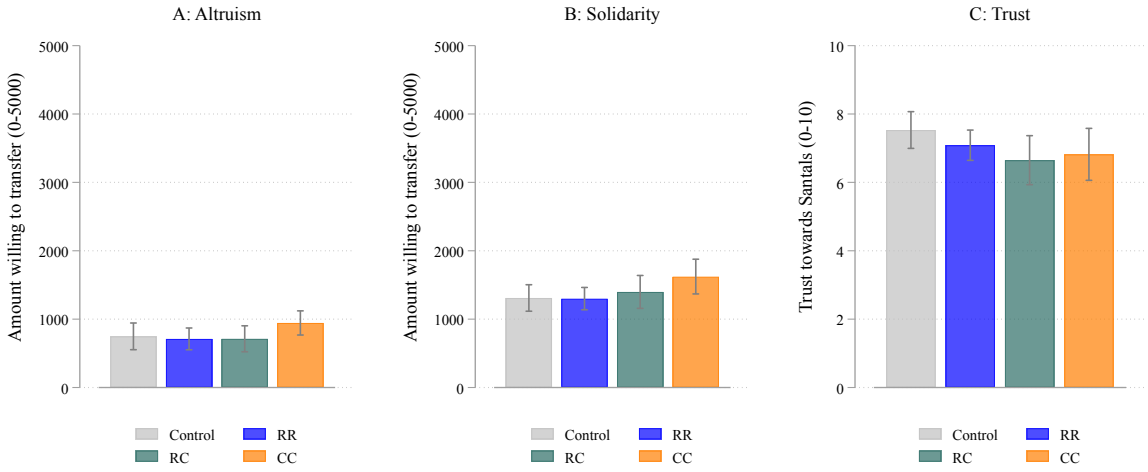
Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. This figure illustrates how our respondents perceived the purpose of the study. After gathering all the data, enumerators asked an open-ended question to the respondents following Chopra et al. (2024), which was: "If you had to guess, what would you say was the purpose of this study?" Enumerators had eight common options to pick from: *Visit/talk*–how often Bengalis and Santals visit or talk to each other, *Lives*–understanding Santals' (Bengalis') lives in the villages, *Correct*–whether documentary film improves Bengalis' attitudes toward Santals, *Giving*–how to donate more money to Santals (Bengalis), *Poor*–how poor people are in these villages, *Help*–understanding how much financial help Santals (Bengalis) need, *Neighbor*–how good are Santals (Bengalis) as neighbors, *Other*–responses that do not fit into any of the other seven categories. Answering option *Correct* would suggest participants could correctly guess the purpose of this study.

Figure A11: Raw differences in prosociality among untreated ethnic majority at endline



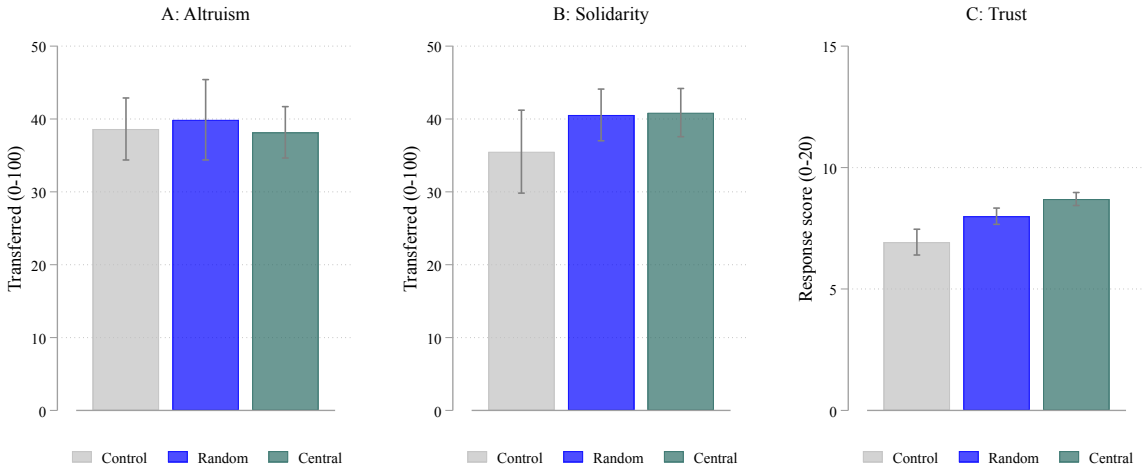
Note: The sample includes untreated Bengalis in the two treatment arm villages and in control villages. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level:
Graph A: Control vs Random: $p = 0.13$; Control vs Central: $p < 0.01$; Random vs Central: $p = 0.12$.
Graph B: Control vs Random: $p = 0.72$; Control vs Central: $p < 0.01$; Random vs Central: $p < 0.01$.
Graph C: Control vs Random: $p = 0.78$; Control vs Central: $p = 0.18$; Random vs Central: $p = 0.41$.

Figure A12: Differences in altruism, solidarity, and trust at baseline



Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. We measured altruism and trust using the survey measure in Falk et al. (2018), which was slightly modified to measure altruism and trust towards non-coethnics. As Falk et al. (2018) do not have a survey measure of solidarity, and given the similarity between altruism and solidarity, we created a survey question closely following the altruism question from Falk et al. (2018). Altruism and solidarity are measured using hypothetical endowments of 5,000 Taka. Trust is measured on a 10-point (verbal) scale. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level:
Graph A: Control vs RR: $p = 0.77$; Control vs RC: $p = 0.80$; Control vs CC: $p = 0.15$; RR vs RC: $p = 0.99$; RR vs CC: $p = 0.06$; RC vs CC: $p = 0.03$.
Graph B: Control vs RR: $p = 0.94$; Control vs RC: $p = 0.57$; Control vs CC: $p = 0.06$; RR vs RC: $p = 0.51$; RR vs CC: $p = 0.04$; RC vs CC: $p = 0.07$.
Graph C: Control vs RR: $p = 0.21$; Control vs RC: $p = 0.06$; Control vs CC: $p = 0.14$; RR vs RC: $p = 0.31$; RR vs CC: $p = 0.55$; RC vs CC: $p = 0.76$.

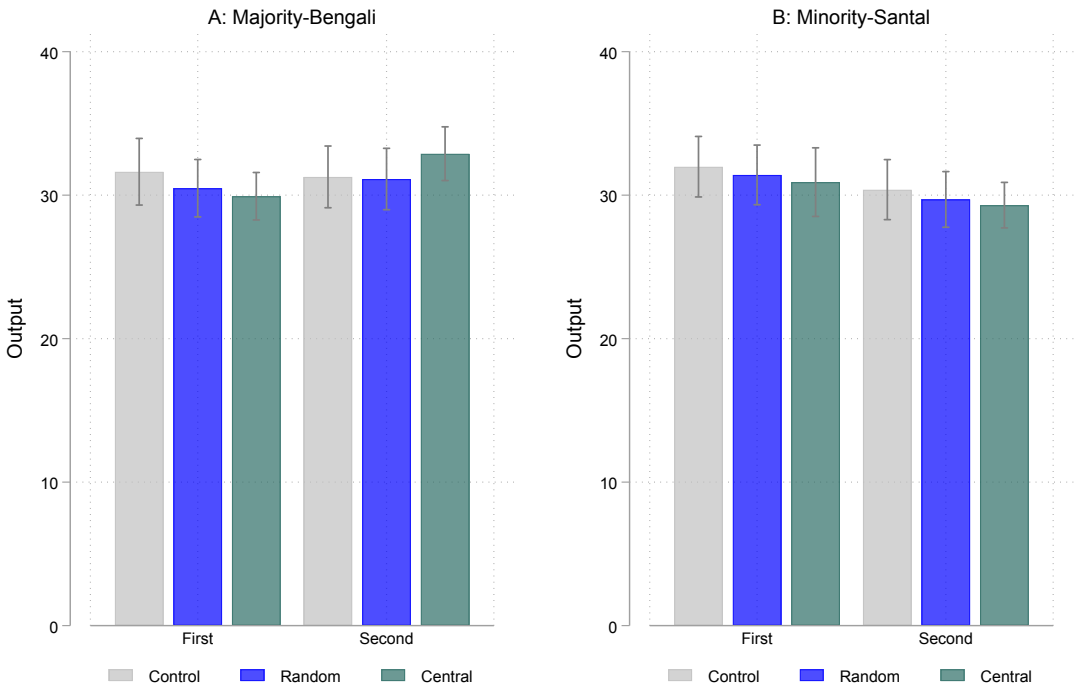
Figure A13: Raw differences in prosociality among untreated ethnic minorities at endline



Note: The sample includes untreated Santals in the two treatment arm villages and in control villages. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level:
Graph A: Control vs Random: $p = 0.72$; Control vs Central: $p = 0.87$; Random vs Central: $p = 0.60$.
Graph B: Control vs Random: $p = 0.14$; Control vs Central: $p = 0.11$; Random vs Central: $p = 0.90$.
Graph C: Control vs Random: $p < 0.01$; Control vs Central: $p < 0.0001$; Random vs Central: $p = 0.02$.

Figure A14: Two casual workers in our field experiment making paper bags

Figure A15: Output as Preparer, by ethnicity



Note: 'First' means a participant started the task as a preparer first but finished as a finisher. 'Second' means a participant finished the task as a preparer but started as a finisher. Bars indicate 95% confidence intervals. All comparisons shown are unconditional (i.e., with no covariate adjustment), and standard errors are clustered at the village level.

A: Majority-Bengali going 'first' as preparer
Control vs Random: $p = 0.47$; Control vs Central: $p = 0.24$; Random vs Central: $p = 0.67$

A: Majority-Bengali going 'second' as preparer
Control vs Random: $p = 0.92$; Control vs Central: $p = 0.27$; Random vs Central: $p = 0.23$

B: Minority-Santal going 'first' as preparer
Control vs Random: $p = 0.71$; Control vs Central: $p = 0.51$; Random vs Central: $p = 0.76$.

B: Minority-Santal going 'second' as preparer
Control vs Random: $p = 0.64$; Control vs Central: $p = 0.42$; Random vs Central: $p = 0.75$.

Table A1: Balance checks of baseline characteristics and outcomes of the ethnic majority Bengali that got treated

VARIABLES	Control (C)		RR		RC		CC		Difference *p*-values					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Randomized comparisons			Comparisons involving CC		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	RC-C	RR-C	RC-RR	CC-C	CC-RR	CC-RC
Age (in years)	39.59	13.24	38.65	12.50	36	13.18	37.73	12.64	0.48	0.78	0.51	0.48	0.62	0.17
Gender (=1 if male)	0.842	0.365	0.879	0.327	0.918	0.275	0.922	0.269	0.38	0.73	0.47	0.36	0.35	0.93
Religion (=1 if Muslim)	0.947	0.224	0.952	0.214	0.986	0.119	0.986	0.119	0.23	0.95	0.11	0.22	0.11	0.87
Household head (=1 if true)	0.785	0.411	0.827	0.379	0.740	0.439	0.776	0.418	0.66	0.04	0.20	0.68	0.23	0.39
Years of education	5.158	4.813	6.148	4.838	6.747	4.769	7.206	4.728	0.08	0.01	0.83	0.01	0.82	0.29
Monthly personal income (in Taka)	8,155	6,741	9,302	6,829	9,727	6,827	10,834	7,873	0.05	0.22	0.46	<0.01	0.06	0.16
Works in farming (=1 if true)	0.576	0.495	0.585	0.493	0.573	0.496	0.537	0.499	0.83	0.96	0.37	0.39	0.61	0.47
Household size	4.808	2.122	4.811	1.905	4.712	1.802	4.744	1.874	0.49	0.83	0.88	0.42	0.81	0.84
Village size (Number of households)	160.8	106.1	137.1	92.17	173.6	88.40	174.8	87.96	0.71	0.66	0.59	0.68	0.64	0.56
Village ethnic diversity	0.362	0.0951	0.383	0.124	0.395	0.114	0.394	0.114	0.56	0.92	0.61	0.51	0.58	0.20
Altruism	748.5	1,089	711.2	1,095	712.8	1,117	944.5	1,373	0.61	0.84	0.51	0.47	0.32	0.06
Solidarity	1,310	1,410	1,300	1,453	1,398	1,591	1,623	1,713	0.81	0.97	0.77	0.23	0.22	0.11
Trust	7.532	2.933	7.087	3.089	6.650	3.862	6.820	4.091	0.07	0.02	0.91	0.16	0.79	0.75
Interactions with Bengali	3.792	0.866	3.905	0.528	3.829	0.712	3.836	0.683	0.41	0.07	0.13	0.37	0.09	0.92
Interactions with Santal	1.864	1.702	1.758	1.744	1.730	1.762	1.925	1.728	0.72	0.27	0.27	0.63	0.14	0.28
Visits to Bengalis	33.45	18.54	32.62	17.72	29.75	19.76	34.87	20.42	0.43	0.78	0.84	0.70	0.99	0.15
Visits to Santals	11.80	15.37	12.25	16.57	11.98	17.00	15.90	17.97	0.94	0.54	0.49	0.12	0.31	0.12
Discriminatory opinion	23.89	7.960	25.87	10.46	29.23	9.763	27.68	9.990	<0.01	0.04	0.34	0.03	0.66	0.22
Stereotypes	13.61	12.02	13.08	11.84	17.60	11.26	19.55	11.43	0.86	0.57	0.85	0.65	0.18	0.33
PHQ-4 score	4.419	5.028	4.237	5.073	3.463	4.248	2.907	3.906	0.44	0.92	0.04	0.14	0.03	0.21
SWB	25.60	7.764	26.27	8.411	26.02	7.561	27.12	7.671	0.63	0.56	0.90	0.18	0.23	0.18
ICC score	1.512	1.232	1.678	1.266	1.612	1.345	1.722	1.331	0.66	0.52	0.90	0.92	0.41	0.45
SDB score	9.132	1.431	9.216	1.515	9.217	1.865	9.249	1.862	0.58	0.23	0.02	0.46	0.33	0.85
Perspective-taking	18.60	3.591	18.68	4.037	18.26	4.803	18.97	4.918	0.92	0.36	0.82	0.42	0.81	0.34
Empathic concern	20.30	4.317	20.93	4.051	20.58	4.838	20.66	5.241	0.32	0.07	0.59	0.28	0.52	0.89
Personal distress	16.68	3.860	16.64	3.735	16.99	3.510	17.08	4.177	0.87	0.67	0.67	0.94	0.57	0.87
Empathy IRI	55.58	9.215	56.21	9.039	55.83	10.23	56.51	12.09	0.68	0.22	0.68	0.51	0.67	0.69
Sample size	**568**		**561**		**281**		**281**							
Joint F-test									**0.16**	**<0.01**	**0.21**	**0.35**	**0.22**	**0.05**

Note: The columns *Control*, *RR*, *RC*, and *CC* represent the four types of samples from three treatment arms, each showing the means and standard deviations of the corresponding variables: *RR* means randomly selected ethnic majorities in the *Random* treatment arm, *RC* means randomly selected ethnic majorities in the *Central* treatment arm, and *CC* means network-central (non-randomly selected) ethnic majorities in the *Central* treatment arm. Difference *p*-values in the first three columns are for randomized comparisons (RC–C, RR–C, RC–RR); the last three columns report comparisons involving CC (CC–C, CC–RR, CC–RC), which are descriptive. All baseline variables are defined in Appendix C.

Table A2: Balance checks of baseline characteristics and outcomes of untreated ethnic majority and minority

VARIABLES	Control		Random		Central		Difference *p*-values		
	Mean	SD	Mean	SD	Mean	SD			
	(1)	(2)	(3)	(4)	(5)	(6)	(5)-(1)	(3)-(1)	(5)-(3)
Panel A: Untreated Ethnic Majority-Bengali									
Age (in years)	39.47	12.63	37.71	12.56	38.36	13.75	0.57	0.29	0.75
Gender (=1 if male)	0.856	0.352	0.893	0.31	0.929	0.258	0.06	0.17	0.60
Religion (=1 if Muslim)	0.964	0.187	0.967	0.18	0.954	0.211	0.82	0.68	0.60
Household head (=1 if true)	0.827	0.379	0.763	0.426	0.749	0.434	0.07	0.19	0.78
Years of education	5.462	4.678	5.87	4.944	6.396	5.029	0.27	0.85	0.16
Monthly personal income (in Taka)	8586	7269	8611	6783	8952	6328	0.27	0.54	0.75
Works in farming (=1 if true)	0.556	0.498	0.493	0.501	0.568	0.496	0.88	0.26	0.32
Household size	4.776	1.863	4.781	1.897	4.646	1.828	0.85	0.74	0.69
Altruism	712.3	1076	973.3	1504	1048	1362	0.80	0.24	0.04
Solidarity	1314	1486	1524	1685	1755	1743	0.34	0.35	0.57
Trust	6.968	3.241	7.385	3.128	6.996	3.684	0.78	0.88	0.54
Interactions with Bengali	3.755	0.904	3.693	1.019	3.441	1.361	0.25	0.99	0.22
Interactions with Santal	1.578	1.663	1.648	1.715	1.743	1.703	0.87	0.36	0.69
Visits to Bengalis	36.69	18.65	32.99	18.82	32.26	20.68	0.20	0.21	0.91
Visits to Santals	11.62	15.37	13.16	17.38	14.48	17.69	0.78	0.66	0.82
Discriminatory opinion	25.08	8.144	24.31	9.978	27.57	8.275	0.20	0.95	0.13
Stereotypes	13.39	10.89	13.3	12.45	19.15	12.94	0.98	0.81	0.68
PHQ-4 score	4.061	5.000	4.856	5.283	4.179	4.440	0.50	0.09	0.21
ICC score	1.361	1.236	1.772	1.325	1.989	1.302	0.92	0.52	0.31
SDB score	9.188	1.48	9.148	1.59	9.114	1.746	0.48	0.45	1.00
Perspective Taking	19.01	3.833	18.99	3.955	18.59	4.912	0.39	0.97	0.71
Empathic Concern	21.0	3.917	21.57	3.958	20.89	4.741	0.34	0.22	0.95
Personal Distress	17.39	3.855	17.42	4.117	17.65	3.249	0.28	0.64	0.10
Empathy IRI	57.40	9.138	57.97	8.835	57.13	10.18	0.68	0.51	0.52
Sample size	**277**		**270**		**280**				
Joint F-test							**0.45**	**0.02**	**0.35**
Panel B: Untreated Ethnic Minority-Santals									
Age (in years)	34.79	13.17	35.3	12.84	34.35	11.95	0.69	0.48	0.56
Gender (=1 if male)	0.892	0.310	0.848	0.36	0.885	0.319	0.21	0.19	0.03
Religion (=1 if Muslim)	0.014	0.119	0.007	0.0859	0.000	0.000	0.45	0.82	0.30
Household head (=1 if true)	0.706	0.456	0.737	0.441	0.667	0.472	0.16	0.99	0.05
Years of education	5.09	4.593	5.207	4.692	5.627	4.655	0.72	0.82	0.55
Monthly personal income (in Taka)	7342	4065	7241	4390	6868	4705	0.84	0.56	0.94
Works in farming (=1 if true)	0.763	0.426	0.696	0.461	0.731	0.444	0.76	0.47	0.54
Household size	4.789	1.457	4.844	1.656	4.663	1.425	0.57	0.56	0.17
Altruism	617.4	910.4	991.7	1301	1031	1374	<0.01	<0.01	0.42
Solidarity	1085	1101	1430	1332	1785	1624	<0.001	<0.01	0.02
Trust	7.282	2.928	6.707	3.483	6.476	3.449	0.96	0.06	0.06
Interactions with Bengali	1.493	1.699	1.574	1.701	1.645	1.759	0.39	1.00	0.35
Interactions with Santal	3.466	1.338	3.1	1.652	3.204	1.566	0.11	0.11	0.71
Visits to Bengalis	15.32	17.49	17.35	18.44	11.91	15.04	0.27	0.67	0.04
Visits to Santals	44.45	19.17	41.89	19.59	35.0	22.0	<0.001	0.35	<0.01
PHQ-4 score	4.376	5.001	5.133	4.967	5.097	4.816	<0.01	0.11	0.14
SWB	25.27	6.725	24.41	6.294	24.06	6.026	0.05	0.94	0.09
SDB score	9.133	1.469	9.222	1.562	9.201	1.477	0.26	0.43	0.54
Sample size	**279**		**270**		**279**				
Joint F-test							**0.14**	**<0.01**	**0.07**

Note: The columns *Control*, *Random*, and *Central* represent the three arms, each showing the means and standard deviations of the corresponding variables; difference *p*-values in the last three columns were computed by regressing the baseline characteristics on the treatment variable with union council fixed effects and standard errors clustered at the village level. All variables are defined in Appendix C.

Table A3: Major deviations from the pre-analysis plan (PAP)

What was pre-specified	**Deviations in the paper**
We said we would measure altruism and solidarity at endline using survey questions	We ended up not measuring these using survey questions because (i) we had incentivized game measures and survey questions could have biased the game measures, (ii) to shorten the endline survey.
Examine the correlation between incentivized and survey measures of preferences	Because of the above reasons, we do not have the survey measure, hence we could not examine this correlation.
Conduct heterogeneity analysis using machine learning following Chernozhukov et al. (2020)	Given that heterogeneity estimates from interaction models were broadly statistically insignificant, we opted not to pursue the Chernozhukov et al. (2020) approach, as any significant results would likely not be robust.
We did not pre-register the qualitative analysis or video content analysis	We did the former to provide more insights about what new things people learned from the documentary and how they processed it. We did the latter as it was suggested by the editor.
Measuring trust using the same 0-10 scale	At endline, trust was measured on a 0-20 scale, using a thermometer scale to increase precision.
For various outcomes, we said we would also code them as binary outcomes (alongside the way we present it)	We did not do it because in many cases, results are statistically insignificant, and showing results using dummy constructed outcomes would not add much value to the result.
Randomize the order of Dictator and Solidarity games	We did not do it because Dictator is a simplified version of the Solidarity game, and we realized during the pilot phase before endline that participants understood the instructions of the Solidarity game better if they played the Dictator game first.
Casual work task would take 2.5 hours per pair	It took 3 hours per pair.
We would do empathy/emotions analysis using self-reported measures	We used instead a richer measure, facial expressions, so we opted not to include the self-reported measures.
The main empirical strategy did not split the *Central* arm into RC and CC	We revised this in the paper by splitting into RC and CC to separate the randomly selected Bengalis from the non-randomly selected Bengalis.
Only select controls using LASSO, with only the baseline measure of outcome as unpenalized control	We included education, income, and baseline outcome as unpenalized controls, as there are differences in education and income across treatments.
Multiple hypothesis testing using two methods and different families	We only do Westfall-Young adjustments. We also define outcomes into several conceptually distinct families (prosociality, self-reported behaviors, interactions/visits, labor productivity, and, by treating village-level complaints separately, given their different units of observation and dataset. We also took a more conservative approach by pooling all behavioral outcomes into one family (prosociality, self-reports, interactions/visits, and productivity) shows very similar qualitative conclusions, with many effects (altruism, trust, and finisher productivity) remaining statistically significant among the *RC* treatment sample.
Heterogeneity using monthly household income, education of participants, age of the participant, gender of the participant, and baseline discriminatory opinions index	We do not present all because, in many cases, there is no heterogeneity, and in other cases, some categories had very few observations, limiting statistical power.
Estimate impacts on productivity and interethnic complaints using equation 1	For productivity, due to random sampling and participant matching between villages, without stratification by union council, in many cases, pairs were from different union councils. Thus, we were unable to use union council FE. For interethnic complaints, we do not select controls using LASSO because we only have 3 control variables.
Analysis of errors as preparer in the casual work experiment	This could not be implemented. We only measured errors on the final products, so we cannot determine whether the error was due to the preparer's mistake or the finisher's mistake.
We only pre-registered social interactions with and visits to non-coethnics	We also collected data on the same towards coethnics and report them in the paper.

Note: The pre-analysis plan of this paper can be downloaded from the AEA RCT Registry: `https://www.socialscienceregistry.org/trials/10730`.

Table A4: Finding treatment effects on pre-registered outcomes in the paper

Pre-specified outcomes	Outcome domain	Table/Figure numbers in the paper	FWER family
A: Main ethnic majority sample			
Altruism, solidarity, trust	Behavior	Figures I–II; Table I	1
Expected solidarity, friendship, water charge, willing to help	Behavior	Figure II; Table I	2
Social interactions	Behavior	Figure II; Table I	3
Stereotypes, opinions	Beliefs	Figure II; Table II	4
Mental health, SWB, IC competence	Other	Figure II; Table A5	5
Photo measures of emotions	-	Figure III; Tables A14–A15	6
Interethnic complaints	Behavior	Table V	7
Productivity as preparer/finisher	Behavior	Figures IV and A15 Tables A26, A27, A28, A29	8
B: Untreated ethnic majority sample			
Altruism, solidarity, trust	Behavior	Figure A11; Table III	1
Expected solidarity, friendship, water charge	Behavior	Table III	2
Social interactions	Behavior	Table III	3
Stereotypes, opinions	Beliefs	Table III (Panel B)	4
Mental health, IC competence	Other	Table A22	5
C: Untreated ethnic minority sample			
Altruism, solidarity, trust	Behavior	Figure A13; Table IV	1
Expected solidarity, friendship, water charge	Behavior	Table IV	2
Social interactions	Behavior	Table IV	3
Mental health, SWB, income, food insecurity, new employment	Other	Table A23	4
Productivity as preparer/finisher	Behavior	Figures IV and A15 Tables A26, A27, A28, A29	5

Note: Interethnic complaints using administrative data are reported under Panel A, but it is a village-level measure and applies to both treated and untreated samples. The pre-analysis plan of this paper can be downloaded from the AEA RCT Registry: `https://www.socialscienceregistry.org/trials/10730`.

Table A5: Treatment effects on other outcomes of treated ethnic majority

	Mental Health	**SWB**	**ICC**
Variables	**(1)**	**(2)**	**(3)**
RR	-0.035	0.026	0.028
	(0.197)	(0.429)	(0.058)
RC	-0.164	0.103	-0.031
	(0.304)	(0.573)	(0.085)
CC	-0.241	0.313	-0.017
	(0.328)	(0.584)	(0.074)
Observations	1,515	1,515	1,507
Control mean	1.613	25.61	2.081
Control SD	2.635	5.381	0.758
RR=RC p-values	0.65	0.88	0.50
RR=CC p-values	0.51	0.60	0.53
RC=CC p-values	0.75	0.67	0.87
RR+RC vs C p-values	0.72	0.91	0.82
FWER *p*-values on RR	0.99	0.99	0.99
FWER *p*-values on RC	0.99	0.99	0.99
FWER *p*-values on CC	0.97	0.97	0.99
RI *p*-values on RR	0.72	0.80	0.68
RI *p*-values on RC	0.63	0.89	0.85
RI *p*-values on CC	0.51	0.54	0.91

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in both treatment arms and the control arm. See the note under Table I for specification details and p-value definitions. All other information is as previously described. Outcomes in the columns are as follows: (1) Mental health is the aggregated PHQ-4 score (0-12), where lower amount corresponds to better mental health; (2) SWB is the aggregate of four subjective well-being indicators from the World Values Survey (0-40), where higher value corresponds to better well-being; (3) ICC is the inter-cultural competence about Santali culture (0-4), where higher value corresponds to higher inter-cultural competence.

Table A6: Behaviors of the treated ethnic majority using Lee bounds

	Unadjusted treatment effects on				Lee (2009) lower bounds		
	RR	RC	CC	Obs.	RR	RC	CC
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Altruism	2.677*	3.252*	3.355*	1,526	1.257	1.408	3.317*
	(1.386)	(1.710)	(1.685)		(1.660)	(1.982)	(1.932)
Solidarity	0.807	1.866	4.137***	1,525	-0.442	0.092	4.104**
	(1.269)	(1.568)	(1.543)		(1.511)	(1.803)	(1.747)
Trust	1.106***	1.345***	3.823***	1,526	0.846***	1.006***	3.816***
	(0.246)	(0.303)	(0.299)		(0.288)	(0.355)	(0.355)
Ex Sol	0.285	1.177	-2.568	1,520	-0.254	0.472	-3.072
	(1.394)	(1.717)	(1.692)		(1.489)	(1.871)	(2.055)
Friends	0.029	0.075	0.016	1,530	-0.043	-0.009	0.014
	(0.109)	(0.134)	(0.132)		(0.116)	(0.150)	(0.134)
Water$	-0.188	-0.321	0.713	446	-0.521	-0.521	-0.521
	(0.835)	(0.880)	(0.837)		(0.521)	(0.521)	(0.521)
Help	0.435***	0.365**	0.588***	1,401	0.397**	0.339*	0.562**
	(0.124)	(0.151)	(0.152)		(0.199)	(0.178)	(0.240)
Int San	0.080	0.001	0.016	1,525	0.014	-0.073	0.008
	(0.062)	(0.077)	(0.076)		(0.073)	(0.091)	(0.085)
Int Ben	-0.019	-0.037**	-0.000	1,525	-0.034***	-0.052***	-0.010
	(0.015)	(0.018)	(0.018)		(0.013)	(0.020)	(0.046)
Visit San	1.793**	0.576	0.833	1,525	1.078	-0.221	0.740
	(0.827)	(1.018)	(1.006)		(0.926)	(1.175)	(1.058)
Visit Ben	-0.336	-0.540	1.318	1,525	-1.805	-2.172	1.117
	(1.097)	(1.351)	(1.335)		(1.371)	(1.729)	(1.555)

Robust standard errors clustered at the village level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Columns (1-3) report unadjusted treatment effects estimated using OLS. Columns (5-7) report the lower bound treatment effects using Lee (2009) bounds. See the table note under Table I for outcome definitions.

Table A7: Disaggregated impacts on social interaction components (treated ethnic majority)

	Interactions with Santals				Interactions with Bengalis			
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
RR	0.007	0.010	0.008	0.002	0.001	-0.003	0.007	0.002
	(0.030)	(0.030)	(0.022)	(0.022)	(0.001)	(0.002)	(0.008)	(0.007)
RC	0.009	0.019	0.002	-0.012	-0.003	0.002	-0.012	-0.012
	(0.032)	(0.033)	(0.029)	(0.030)	(0.003)	(0.002)	(0.011)	(0.011)
CC	0.024	0.038	-0.003	-0.016	0.000	0.002	0.004	0.005
	(0.034)	(0.033)	(0.031)	(0.033)	(0.001)	(0.002)	(0.010)	(0.010)
Observations	1,516	1,518	1,517	1,515	1,514	1,515	1,516	1,514
Control mean	0.818	0.816	0.0746	0.0766	1.000	1.000	0.992	0.992
Control SD	0.386	0.387	0.263	0.266	0.000	0.000	0.0872	0.0872
RR=RC p-values	0.93	0.77	0.83	0.60	0.25	0.20	0.05	0.20
RR=CC p-values	0.62	0.41	0.71	0.53	0.61	0.22	0.77	0.80
RC=CC p-values	0.64	0.54	0.87	0.88	0.31	0.23	0.18	0.18
RR+RC vs C p-values	0.78	0.65	0.77	0.92	0.59	0.26	0.80	0.78

Robust standard errors clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. Outcomes in the columns are as follows: (1) =1 if offer Santals food when they visit my house; (2) =1 if offer Santals chair/seat when they visit my house; (3) =1 if invite Santals during festivals; (4) =1 if Santals invite me during festivals; (5) =1 if offer Bengalis food when they visit my house; (6) =1 if offer Bengalis chair/seat when they visit my house; (7) =1 if invite Bengalis during festivals; (8) =1 if Bengalis invite me during festivals.

Table A8: Heterogeneity in prosociality of treated ethnic majority, by baseline stereotypes

VARIABLES	(1) Altruism	(2) Solidarity	(3) Trust
Treatment RR	-1.067	2.755	0.625*
	(3.131)	(2.141)	(0.359)
Treatment RC	5.547*	7.862***	1.399**
	(3.230)	(2.584)	(0.581)
Treatment CC	4.484	9.025***	3.127***
	(3.953)	(2.938)	(0.633)
Stereotypes score	-0.029	0.152**	-0.022
	(0.092)	(0.062)	(0.015)
Stereotype score × RR	0.256**	-0.004	0.015
	(0.122)	(0.085)	(0.021)
Stereotype score × RC	0.032	-0.282**	0.011
	(0.150)	(0.131)	(0.026)
Stereotype score × CC	0.123	-0.215	0.049*
	(0.172)	(0.132)	(0.028)
Constant	38.696***	33.568***	7.913***
	(3.360)	(3.070)	(0.782)
Observations	1,505	1,504	1,502

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. We define the stereotype measure in section 3.2.1 of the paper.

Table A9: Heterogeneity in beliefs and interactions of treated ethnic majority, by baseline stereotypes

VARIABLES	(1) Stereotype	(2) Discr Opinions	(3) Interact-Santals	(4) Interact-Bengalis	(5) Visit-Santals	(6) Visit-Bengalis	(7) ICC score
Treatment RR	0.003	-0.082	0.003	0.000	0.019	0.115	-0.067
	(0.107)	(0.105)	(0.125)	(0.124)	(0.095)	(0.104)	(0.104)
Treatment RC	0.211*	0.279*	0.056	0.241	-0.012	0.226	-0.185
	(0.128)	(0.147)	(0.146)	(0.162)	(0.137)	(0.145)	(0.158)
Treatment CC	0.022	-0.035	-0.038	-0.046	0.128	0.295**	-0.038
	(0.124)	(0.155)	(0.153)	(0.242)	(0.129)	(0.129)	(0.152)
Stereotypes score	0.000	-0.001	-0.002	-0.003	-0.006	0.003	-0.002
	(0.005)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
Stereotype score × RR	-0.002	0.000	0.002	0.004	0.004	-0.007	0.008
	(0.006)	(0.007)	(0.006)	(0.004)	(0.005)	(0.006)	(0.005)
Stereotype score × RC	-0.015**	-0.014**	-0.003	-0.023*	0.004	-0.015**	0.009
	(0.007)	(0.007)	(0.008)	(0.013)	(0.007)	(0.007)	(0.007)
Stereotype score × CC	-0.001	0.000	0.004	0.006	-0.001	-0.012**	0.002
	(0.006)	(0.007)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)
Constant	-0.057	-0.088	-0.010	-0.049	0.244	-0.048	0.101
	(0.159)	(0.179)	(0.153)	(0.152)	(0.151)	(0.156)	(0.134)
Observations	1,476	1,476	1,477	1,477	1,477	1,476	1,477

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. We define the stereotype measure in section 3.2.2 in the paper.

Table A10: Heterogeneity in prosociality of treated ethnic majority, by baseline discriminatory opinions

VARIABLES	(1) Altruism	(2) Solidarity	(3) Trust
Treatment RR	2.324	11.929***	0.565
	(4.436)	(3.900)	(0.695)
Treatment RC	11.003**	11.042**	0.967
	(5.496)	(4.343)	(0.945)
Treatment CC	7.390	17.813***	2.820***
	(4.618)	(4.329)	(0.924)
Discriminatory opinion scores	0.119	0.345***	-0.020
	(0.124)	(0.094)	(0.022)
Discr Opinions × RR	-0.009	-0.397***	0.014
	(0.149)	(0.127)	(0.028)
Discr Opinions × RC	-0.195	-0.345**	0.025
	(0.183)	(0.139)	(0.035)
Discr Opinions × CC	-0.042	-0.510***	0.047
	(0.164)	(0.137)	(0.035)
Constant	35.740***	27.100***	7.781***
	(4.249)	(3.774)	(0.848)
Observations	1,513	1,512	1,510

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. We define the discriminatory opinions measure in section 3.2.1 in the paper.

Table A11: Heterogeneity in beliefs and interactions of treated ethnic majority, by baseline discriminatory opinions

VARIABLES	(1) Stereotype	(2) Discr Opinions	(3) Interact-Santals	(4) Interact-Bengalis	(5) Visit-Santals	(6) Visit-Bengalis	(7) ICC score
Treatment RR	-0.232	-0.456**	-0.010	-0.370	0.032	0.059	0.014
	(0.178)	(0.190)	(0.208)	(0.249)	(0.202)	(0.187)	(0.196)
Treatment RC	-0.155	0.141	-0.250	0.174	-0.408*	-0.005	-0.083
	(0.230)	(0.230)	(0.216)	(0.325)	(0.246)	(0.233)	(0.257)
Treatment CC	-0.156	0.049	-0.060	0.315	0.147	0.221	-0.147
	(0.225)	(0.251)	(0.231)	(0.230)	(0.227)	(0.205)	(0.228)
Discriminatory opinion scores	-0.007	-0.007	-0.001	-0.004	-0.008	0.000	0.005
	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)	(0.005)	(0.006)
Discr Opinions × RR	0.008	0.015**	0.002	0.016*	0.003	-0.002	0.001
	(0.006)	(0.007)	(0.007)	(0.009)	(0.008)	(0.007)	(0.008)
Discr Opinions × RC	0.005	-0.002	0.009	-0.010	0.017*	-0.001	0.001
	(0.008)	(0.008)	(0.008)	(0.013)	(0.009)	(0.008)	(0.009)
Discr Opinions × CC	0.007	-0.001	0.004	-0.008	-0.001	-0.006	0.004
	(0.008)	(0.009)	(0.009)	(0.011)	(0.008)	(0.008)	(0.009)
Constant	0.149	0.069	-0.010	-0.019	0.354*	-0.036	-0.083
	(0.179)	(0.181)	(0.191)	(0.225)	(0.184)	(0.179)	(0.170)
Observations	1,476	1,485	1,486	1,486	1,486	1,485	1,478

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. We define the discriminatory opinions measure in section 3.2.2 in the paper. In column 2, education and income were not included as unpenalized controls because of an error.

Table A12: Heterogeneity in prosociality of treated ethnic majority, by ethnic composition of village

VARIABLES	(1) Altruism	(2) Solidarity	(3) Trust
Treatment RR	11.822**	4.637	0.927
	(5.813)	(4.192)	(0.898)
Treatment RC	8.984*	10.901*	1.356
	(4.802)	(6.340)	(1.117)
Treatment CC	21.315***	18.986***	5.589***
	(7.012)	(4.784)	(1.070)
Proportion of Santals in a village	22.897***	7.549	0.719
	(8.746)	(8.362)	(1.476)
% of Santals in Village × RR	-25.104*	-5.958	-0.119
	(14.426)	(10.762)	(2.138)
% of Santals in Village × RC	-9.147	-22.109	0.600
	(11.547)	(17.338)	(2.487)
% of Santals in Village × CC	-39.013**	-36.963***	-3.950
	(17.626)	(11.685)	(2.704)
Constant	30.449***	34.210***	6.942***
	(4.166)	(3.775)	(1.042)
Observations	1,519	1,518	1,516

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. Proportion of Santals in a village = number of Santals households divided by the total number of households, so this value is between 0 and 1.

Table A13: Heterogeneity in beliefs and interactions of treated ethnic majority, by ethnic composition of village

VARIABLES	(1) Stereotype	(2) Discr Opinions	(3) Interact-Santals	(4) Interact-Bengalis	(5) Visit-Santals	(6) Visit-Bengalis	(7) ICC score
Treatment RR	-0.026	0.057	-0.348	0.021	-0.013	-0.108	0.293
	(0.247)	(0.274)	(0.214)	(0.266)	(0.199)	(0.270)	(0.215)
Treatment RC	-0.009	0.101	-0.285	0.114	0.030	-0.423	0.411
	(0.303)	(0.367)	(0.260)	(0.299)	(0.259)	(0.327)	(0.317)
Treatment CC	-0.159	-0.071	0.216	0.001	0.105	-0.198	0.087
	(0.278)	(0.312)	(0.281)	(0.203)	(0.250)	(0.282)	(0.211)
Proportion of Santals in a village	-0.007	0.089	-0.445	-0.091	-0.276	-0.929*	0.888**
	(0.517)	(0.635)	(0.402)	(0.313)	(0.460)	(0.542)	(0.352)
% of Santals in Village × RR	-0.022	-0.372	1.022**	0.050	0.250	0.296	-0.665
	(0.634)	(0.695)	(0.517)	(0.585)	(0.512)	(0.690)	(0.589)
% of Santals in Village × RC	-0.066	-0.178	0.790	-0.706	0.076	1.004	-1.152
	(0.776)	(0.942)	(0.651)	(0.813)	(0.659)	(0.826)	(0.743)
% of Santals in Village × CC	0.444	0.098	-0.401	0.132	-0.050	0.689	-0.298
	(0.683)	(0.767)	(0.718)	(0.383)	(0.668)	(0.671)	(0.513)
Constant	-0.038	-0.193	0.108	-0.069	0.236	0.328	-0.278*
	(0.234)	(0.281)	(0.185)	(0.182)	(0.195)	(0.239)	(0.164)
Observations	1,479	1,488	1,492	1,492	1,492	1,491	1,484

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. The proportion of Santals in a village = number of Santals households divided by the total number of households, so this value is between 0 and 1.

Table A14: Treatment effects on emotions using facial expressions of treated ethnic majority

	Anger	Disgust	Fear	Happy	Sad	Surprise
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Random in Random (RR)	0.521	0.078	0.022	0.502	2.186	0.015
	(0.964)	(0.208)	(0.037)	(0.829)	(1.530)	(0.040)
Random in Central (RC)	-0.676	0.169	0.055	1.432	3.796*	0.043
	(1.181)	(0.256)	(0.063)	(1.130)	(2.084)	(0.110)
Central in Central (CC)	-1.111	0.588	-0.037	2.279**	6.293***	-0.007
	(1.021)	(0.471)	(0.046)	(1.123)	(2.353)	(0.059)
RR=RC p-values	0.35	0.70	0.59	0.40	0.42	0.80
RR=CC p-values	0.14	0.26	0.17	0.08	0.08	0.70
RC=CC p-values	0.66	0.12	0.06	0.47	0.25	0.57
Observations	1,472	1,472	1,472	1,472	1,472	1,472
Control mean	5.11	1.04	0.35	4.86	18.68	0.88
	[13.02]	[2.60]	[0.78]	[10.91]	[23.42]	[1.17]

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. The sum of all six emotions reported above, along with the neutral emotion (i.e., no emotions), equals 100. This means that each individual's emotions were scored between 0-100. A higher value indicates a stronger presence of that particular emotion.

Table A15: Treatment effects on emotions using facial expressions of treated ethnic majority (random sample only)

	Anger	Disgust	Fear	Happy	Sad	Surprise
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Pooled treatment (RR+RC)	0.286	0.079	0.026	0.572	2.599*	0.019
	(0.908)	(0.180)	(0.039)	(0.787)	(1.361)	(0.046)
Observations	1,222	1,222	1,222	1,222	1,222	1,222
Control mean	5.11	1.04	0.35	4.86	18.68	0.88
	[13.02]	[2.60]	[0.78]	[10.91]	[23.42]	[1.17]

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis that are randomly selected in the two treatment arm villages and in control villages. This means that network-central people have been dropped from the analysis. Pooled treatment (RR+RC) is a dummy that equals to 1 if randomly selected Bengalis were either in the RC or RR group, or equals to 0 otherwise. The sum of all six emotions reported above, along with the neutral emotion (i.e., no emotions), equals 100. This means that each individual's emotions were scored between 0-100. A higher value indicates a stronger presence of that particular emotion.

Table A16: Heterogeneity in prosociality of treated ethnic majority, by baseline empathic tendencies

VARIABLES	(1) **Altruism**	(2) **Altruism**	(3) **Altruism**	(4) **Solidarity**	(5) **Solidarity**	(6) **Solidarity**	(7) **Trust**	(8) **Trust**	(9) **Trust**
Treatment RR	3.153	2.501	16.791***	9.379	-5.643	7.362	-0.832	0.241	0.243
	(8.011)	(7.240)	(6.282)	(6.040)	(7.415)	(5.477)	(1.407)	(1.346)	(1.317)
Treatment RC	10.600	15.569**	19.435**	12.055*	4.528	11.906*	1.252	3.105**	2.161
	(7.923)	(6.479)	(8.030)	(6.209)	(6.757)	(6.539)	(1.475)	(1.342)	(1.321)
Treatment CC	-1.372	1.309	17.082*	12.626*	-13.370*	6.428	5.068***	5.469***	3.207**
	(8.968)	(8.133)	(9.222)	(7.069)	(6.868)	(7.448)	(1.162)	(1.294)	(1.302)
Perspective taking (PT) score	0.062			0.410*			-0.017		
	(0.275)			(0.217)			(0.055)		
Perspective taking × RR	-0.036			-0.371			0.090		
	(0.392)			(0.300)			(0.076)		
Perspective taking × RC	-0.263			-0.479			0.021		
	(0.376)			(0.321)			(0.082)		
Perspective taking × CC	0.406			-0.376			-0.050		
	(0.425)			(0.356)			(0.065)		
Empathic concern (EC) score		0.101			-0.402**			0.033	
		(0.192)			(0.192)			(0.037)	
Empathic concern × RR		-0.003			0.405			0.028	
		(0.351)			(0.352)			(0.064)	
Empathic concern × RC		-0.465			-0.073			-0.072	
		(0.302)			(0.318)			(0.061)	
Empathic concern × CC		0.255			0.913***			-0.067	
		(0.370)			(0.327)			(0.059)	
Personal distress (PD) score			0.335			0.132			-0.002
			(0.255)			(0.228)			(0.048)
Personal distress × RR			-0.842**			-0.295			0.037
			(0.366)			(0.302)			(0.079)
Personal distress × RC			-0.783*			-0.550			-0.034
			(0.448)			(0.372)			(0.077)
Personal distress × CC			-0.610			-0.107			0.047
			(0.505)			(0.434)			(0.076)
Constant	37.537***	36.958***	32.833***	28.202***	44.496***	33.887***	7.654***	6.663***	7.209***
	(6.578)	(4.712)	(5.414)	(5.326)	(5.060)	(5.221)	(1.081)	(0.985)	(0.958)
Observations	1,507	1,507	1,507	1,506	1,506	1,506	1,504	1,504	1,504

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. Outcomes in all columns are the prosociality index. We measure the three types of empathic tendencies of participants using the Interpersonal Reactivity Index by Davis (1983): perspective-taking or PT (the tendency to spontaneously adopt the psychological view of others in everyday life), empathic concern or EC (the tendency to experience feelings of sympathy or compassion for unfortunate others), and personal distress or PD (the tendency to experience distress or discomfort in response to extreme distress in others). All three types of empathic tendencies are continuous variables, where a higher value corresponds to higher empathy of that type.

Table A17: Experimenter demand effects check: Effects on behaviors of treated ethnic majority

	Prosociality			**Self-reports**				**Interactions & visits**			
	Altruism	**Solidarity**	**Trust**	**Ex Sol**	**Friends**	**Water$**	**Help**	**Int San**	**Int Ben**	**Visit San**	**Visit Ben**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
RR	2.541	1.875	0.906***	0.560	-0.087	-0.875*	0.399***	0.051	0.007	1.445*	0.670
	(2.132)	(1.550)	(0.317)	(1.731)	(0.124)	(0.519)	(0.149)	(0.078)	(0.016)	(0.842)	(1.330)
RC	6.189***	2.341	1.657***	3.331	0.137	-1.689	0.234	0.051	-0.027	0.675	-0.202
	(2.048)	(1.943)	(0.463)	(2.531)	(0.164)	(1.068)	(0.188)	(0.080)	(0.023)	(1.020)	(1.764)
CC	6.374***	4.700**	4.053***	-0.435	0.058	-0.606	0.449**	0.062	0.011	1.176	1.690
	(2.306)	(1.833)	(0.437)	(2.336)	(0.156)	(0.948)	(0.205)	(0.087)	(0.020)	(0.973)	(1.692)
Observations	1,483	1,483	1,480	1,478	1,486	438	1,363	1,482	1,482	1,482	1,481
Number of groups	0	0	0	0	0	0	0	0	0	0	0
Control mean	40.61	35.71	7.417	29.98	1.866	0.532	6.558	1.782	3.984	19.30	48.49
Control SD	23.82	21.97	3.794	22.59	1.715	5.157	2.053	0.927	0.165	12.81	17.91
RR=RC p-values	0.10	0.80	0.09	0.27	0.17	0.27	0.31	1.00	0.14	0.41	0.62
RR=CC p-values	0.13	0.10	<0.01	0.67	0.35	0.71	0.78	0.90	0.84	0.75	0.54
RC=CC p-values	0.93	0.20	<0.01	0.07	0.59	0.26	0.17	0.90	0.14	0.60	0.20
RR+RC vs C p-value	0.06	0.17	<0.001	0.41	0.85	0.10	0.01	0.46	0.86	0.12	0.73

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: Same as Table I, but dropped the 2% sample (33 people) that correctly guessed the hypothesis of this study.

Table A18: Conservative experimenter demand effects check: Effects on behaviors of treated ethnic majority

	Prosociality			Self-reports				Interactions & visits			
	Altruism	Solidarity	Trust	Ex Sol	Friends	Water$	Help	Int San	Int Ben	Visit San	Visit Ben
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
RR	1.046	4.708*	1.141***	1.297	-0.248	-1.312	0.388	-0.032	0.012	1.680	1.638
	(2.270)	(2.669)	(0.419)	(2.631)	(0.199)	(0.855)	(0.239)	(0.114)	(0.020)	(1.270)	(1.651)
RC	3.507	1.567	1.234**	3.214	-0.001	-2.165	0.298	0.078	-0.018	1.230	0.997
	(2.825)	(2.736)	(0.548)	(2.739)	(0.301)	(1.613)	(0.304)	(0.121)	(0.029)	(1.477)	(2.128)
CC	4.740	4.621	3.680***	0.079	-0.225	-0.318	0.574	0.025	0.037	0.471	4.071*
	(3.419)	(2.820)	(0.639)	(2.734)	(0.281)	(1.507)	(0.356)	(0.136)	(0.038)	(1.555)	(2.393)
Observations	728	728	727	726	731	214	679	729	730	729	730
Control mean	42.42	35.30	7.193	28.92	1.959	1.087	6.525	1.793	3.981	18.74	47.77
Control SD	25.24	22.70	3.785	21.84	1.854	7.372	2.102	0.988	0.182	13.25	18.04
RR=RC p-values	0.42	0.21	0.87	0.55	0.41	0.42	0.73	0.33	0.29	0.76	0.78
RR=CC p-values	0.32	0.98	<0.001	0.69	0.94	0.51	0.56	0.66	0.53	0.42	0.33
RC=CC p-values	0.69	0.28	<0.001	0.22	0.38	0.32	0.19	0.69	0.19	0.67	0.19
RR+RC vs C p-value	0.38	0.13	<0.01	0.40	0.39	0.14	0.12	0.98	0.88	0.18	0.33

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: Same as Table I, with the exclusion of the sample of participants who either guessed the true hypothesis of the study or indicated that the study is about Santal peoples' lives.

Table A19: Heterogeneity in prosociality of treated ethnic majority, by baseline SDB score

VARIABLES	(1) Altruism	(2) Solidarity	(3) Trust
Treatment RR	4.079	1.220	-0.246
	(8.547)	(8.156)	(1.363)
Treatment RC	-3.732	-2.265	0.304
	(8.612)	(7.850)	(1.619)
Treatment CC	-12.353	-5.385	2.137
	(8.489)	(7.952)	(1.880)
Social desirability bias score (0-13)	-0.810	-0.693	-0.116
	(0.626)	(0.636)	(0.098)
SDB score × RR	-0.171	0.131	0.121
	(0.849)	(0.877)	(0.140)
SDB score × RC	1.043	0.556	0.141
	(0.859)	(0.838)	(0.162)
SDB score × CC	2.036**	1.143	0.208
	(0.860)	(0.862)	(0.191)
Constant	46.631***	42.772***	8.288***
	(6.514)	(6.204)	(1.174)
Observations	1,516	1,515	1,513

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. SDB Score is a score between 0-13, where a higher score means someone has a higher tendency to give socially desirable responses in surveys (Dhar et al., 2022).

Table A20: Heterogeneity in beliefs and interactions of treated ethnic majority, by baseline SDB score

VARIABLES	(1) Stereotype	(2) Discr Opinions	(3) Interact-San	(4) Interact-Ben	(5) Visit-San	(6) Visit-Ben	(7) ICC score
Treatment RR	-0.638	-0.710*	-0.238	0.154	-0.999**	-0.339	-0.228
	(0.418)	(0.379)	(0.436)	(0.322)	(0.407)	(0.470)	(0.437)
Treatment RC	-0.120	-0.346	0.341	-0.563	-0.063	0.372	-0.173
	(0.457)	(0.426)	(0.415)	(0.666)	(0.411)	(0.411)	(0.424)
Treatment CC	0.138	-0.500	-0.260	-0.334	0.026	0.485	-0.759**
	(0.415)	(0.392)	(0.420)	(0.310)	(0.415)	(0.398)	(0.375)
Social desirability bias score (0-13)	-0.044*	-0.050*	-0.035	-0.037	-0.054	-0.001	-0.036
	(0.026)	(0.029)	(0.030)	(0.027)	(0.033)	(0.033)	(0.030)
SDB score × RR	0.065	0.068	0.028	-0.012	0.117***	0.038	0.030
	(0.046)	(0.041)	(0.048)	(0.037)	(0.044)	(0.051)	(0.048)
SDB score × RC	0.009	0.041	-0.035	0.044	0.014	-0.043	0.015
	(0.048)	(0.043)	(0.043)	(0.071)	(0.045)	(0.045)	(0.045)
SDB score × CC	-0.013	0.051	0.034	0.043	0.008	-0.044	0.080**
	(0.043)	(0.040)	(0.044)	(0.035)	(0.044)	(0.042)	(0.040)
Constant	0.344	0.298	0.241	0.218	0.599*	-0.047	0.390
	(0.262)	(0.301)	(0.312)	(0.248)	(0.323)	(0.319)	(0.288)
Observations	1,476	1,485	1,489	1,489	1,489	1,488	1,481

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes treated Bengalis in the two treatment arm villages and in control villages. SDB Score is a score between 0-13, where a higher score means someone has a higher tendency to give socially desirable responses in surveys (Dhar et al., 2022).

Table A21: Correlation between outcomes of untreated ethnic majority and their social proximity to network central ethnic majority people

VARIABLES	(1) Altruism	(2) Solidarity	(3) Trust	(4) Stereotype	(5) Discr Opinions	(6) Altruism	(7) Solidarity	(8) Trust	(9) Stereotype	(10) Discr Opinions
Average distance to central	0.189	1.064	-0.086	-0.058***	0.066*					
	(0.682)	(0.740)	(0.122)	(0.019)	(0.040)					
Visits to central						-0.189	0.580**	-0.068***	-0.011	0.063***
						(0.234)	(0.229)	(0.023)	(0.007)	(0.016)
Constant	26.776***	47.220***	9.200***	0.141	0.485*	33.234***	41.996***	8.483***	0.220**	0.037
	(4.732)	(4.765)	(0.769)	(0.125)	(0.267)	(3.064)	(4.117)	(0.416)	(0.099)	(0.183)
Observations	263	263	261	263	263	800	800	797	780	784

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Average distance to central: The average amount of time it takes (in minutes) to walk to the homes of the seven network-central people in a village. Visits to central: The total number of times someone visited the homes of all seven network-central people in their village.

Table A22: Treatment effects on other outcomes of untreated ethnic majority

	Mental Health	**ICC**
Variables	**(1)**	**(2)**
Random (R)	-0.356	-0.007
	(0.381)	(0.074)
Central (C)	-0.372	0.128
	(0.414)	(0.103)
Observations	800	795
Control mean	4.231	2.572
Control SD	2.693	0.648
R=C p-value	0.97	0.24
FWER p-values on R	0.65	0.87
FWER p-values on C	0.65	0.53

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes untreated Bengalis in both treatment arms and the control arm. See the note under Table I for specification details and p-value definitions, and Table A5 for outcome definitions. All other information is as previously described.

Table A23: Treatment effects on other outcomes of untreated ethnic minority

	Mental Health	SWB	Income	Food Insecurity	New Employment
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Random (R)	-0.177	6.159***	-379.902	-0.383	0.000
	(0.532)	(1.103)	(385.329)	(0.305)	(0.085)
Central (C)	-0.955	8.511***	1,021.722	-0.731*	-0.014
	(0.636)	(1.049)	(745.475)	(0.375)	(0.092)
Observations	721	721	721	721	722
Control mean	4.030	16.65	8264	2.652	0.408
Control SD	1.603	4.125	1943	1.417	0.493
R=C p-value	0.27	0.02	0.11	0.28	0.80
FWER p-values on R	0.96	<0.001	0.41	0.34	0.99
FWER p-values on C	0.21	<0.001	0.21	0.07	0.96

Robust standard errors clustered at village-level
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample includes untreated Santals in both treatment arms and the control arm. See the note under Table I for specification details and p-value definitions. All other information is as previously described. Outcomes in the columns are as follows: (1) Mental health is the aggregated PHQ-4 score (0-12), where lower amount corresponds to better mental health; (2) SWB is the aggregate of four subjective well-being indicators from the World Values Survey (0-40), where higher value corresponds to better well-being; (3) Income is the monthly household income; (4) Food insecurity is an index (where a lower value indicates a more favorable outcome), and (5) New Employment is a dummy variable that is equal to 1 if the person started a new job recently and 0 otherwise.

Table A24: Balance checks of baseline characteristics and outcomes of **ethnic majority** who participated in the casual work task

VARIABLES	**Control**		**Random**		**Central**				
	Mean	SD	Mean	SD	Mean	SD	**Difference *p*-values**		
	(1)	(2)	(3)	(4)	(5)	(6)	(5)-(1)	(3)-(1)	(5)-(3)
Age (in years)	40.25	12.84	37.24	12.47	36.27	13.70	0.43	0.57	0.80
Religion (=1 if Muslim)	0.933	0.250	0.942	0.235	0.983	0.129	0.21	0.99	0.25
Household head (=1 if true)	0.808	0.395	0.775	0.419	0.717	0.453	0.51	0.67	0.73
Years of education	5.408	5.156	6.508	4.429	7.192	4.577	0.08	0.48	0.30
Monthly personal income (in Taka)	8,996	5,768	10,546	5,535	10,200	5,379	0.08	0.01	0.87
Works in farming (=1 if true)	0.583	0.495	0.633	0.484	0.567	0.498	0.50	0.52	0.13
Household size	5.017	2.264	5.050	2.024	4.833	1.839	0.96	0.79	0.93
Village size (no. of households)	162.9	100.1	132.3	90.56	166.4	89.21	0.70	0.23	0.47
Village ethnic diversity	0.351	0.0880	0.391	0.130	0.399	0.109	0.44	0.80	0.46
Altruism	733.8	1,158	755	1,148	676.2	1,099	0.50	0.28	0.63
Solidarity	1,368	1,482	1,416	1,654	1,580	1,723	0.21	0.35	0.60
Trust	7.242	3.117	7.142	3.247	7.000	3.735	0.55	0.43	0.55
Interactions with Bengali	3.908	0.550	3.958	0.201	3.883	0.611	0.75	0.36	0.84
Interactions with Santal	1.933	1.743	1.775	1.732	1.525	1.763	0.25	0.68	0.73
Visits to Bengalis	34.58	18.35	33.99	17.22	28.77	19.36	0.15	0.69	0.40
Visits to Santals	12.16	15.70	14.07	19.07	8.975	13.27	0.42	0.84	0.83
Discriminatory opinion	23.77	7.899	25.78	9.993	28.34	10.17	<0.01	0.15	0.50
Stereotypes	13.08	11.70	12.59	11.08	15.76	11.36	0.26	0.47	0.42
MH score	4.417	5.180	4.458	5.056	2.892	3.856	0.11	0.88	0.06
SWB	25.37	8.052	25.88	8.193	27.37	7.199	0.15	0.61	0.24
ICC score	1.517	1.167	1.754	1.205	1.725	1.334	0.12	0.27	0.42
SDB score	9.108	1.413	9.258	1.452	9.283	1.802	0.27	0.52	0.11
Perspective Taking	19.06	3.513	19.63	4.159	18.46	4.775	0.85	0.51	0.80
Empathic Concern	20.82	3.716	21.54	4.021	20.82	4.836	0.78	0.30	0.95
Personal Distress	16.77	3.823	16.86	3.694	16.98	3.489	0.73	0.85	0.58
Empathy IRI	56.65	8.422	58.03	8.970	56.26	10.24	0.90	0.44	0.98
Sample size	**120**		**120**		**120**				
Joint F-test							**0.28**	**0.32**	**0.24**

Note: The columns *Control*, *Random*, and *Central* represent the three arms, each showing the means and standard deviations of the corresponding variables; difference *p*-values in the last three columns were computed by regressing the baseline characteristics on the treatment variable with union council fixed effects and standard errors clustered at the village level. All variables are defined in Appendix C of the paper.

Table A25: Balance checks of baseline characteristics and outcomes of **ethnic minority** who participated in the casual work task

VARIABLES	Control		Random		Central				
	Mean	SD	Mean	SD	Mean	SD	**Difference *p*-values**		
	(1)	(2)	(3)	(4)	(5)	(6)	(5)-(1)	(3)-(1)	(5)-(3)
Age (in years)	35.76	13.12	36.53	13.14	35.23	12.66	0.83	0.85	0.70
Religion (=1 if Muslim)	0.017	0.129	0.008	0.091	0.00	0.00	0.32	0.31	<0.01
Household head (=1 if true)	0.742	0.440	0.775	0.419	0.683	0.467	0.25	0.98	0.16
Years of education	5.050	4.381	5.642	4.738	5.533	4.634	0.66	0.15	0.56
Monthly personal income (in Taka)	8,246	3,682	7,771	4,586	7,532	4,527	0.42	0.62	0.94
Works in farming (=1 if true)	0.808	0.395	0.675	0.470	0.792	0.408	0.97	0.03	0.02
Household size	4.750	1.361	4.917	1.617	4.700	1.504	0.69	0.20	0.25
Village size (no. of households)	165.9	106.9	135.6	90.12	171.2	88.53	0.34	0.84	0.35
Village ethnic diversity	0.358	0.0975	0.387	0.119	0.391	0.107	0.47	0.85	0.53
Altruism	604.6	828.1	976.3	1,231	966.0	1,313	0.04	0.01	0.99
Solidarity	997.5	1,021	1,408	1,240	1,778	1,585	<0.01	0.01	0.45
Trust	7.218	3.017	6.650	3.392	6.610	3.419	0.68	0.03	0.08
Interactions with Bengali	1.622	1.727	1.617	1.740	1.708	1.812	0.36	0.84	0.62
Interactions with Santal	3.483	1.316	3.108	1.664	3.175	1.586	0.17	0.24	0.68
Visits to Bengalis	16.91	17.52	14.52	16.65	11.77	14.66	0.27	0.23	0.64
Visits to Santals	47.64	17.60	41.67	19.84	32.97	22.31	<0.01	0.12	0.01
MH score	3.692	4.761	5.175	4.862	5.217	4.767	0.02	0.03	0.47
SWB	26.01	6.154	23.89	6.012	23.47	5.648	0.01	0.50	0.02
SDB score	9.200	1.363	9.017	1.725	9.217	1.567	0.95	0.19	0.59
Sample size	**120**		**120**		**120**				
Joint F-test							**0.24**	**0.01**	**0.11**

Note: The columns *Control*, *Random*, and *Central* represent the three arms, each showing the means and standard deviations of the corresponding variables; difference *p*-values in the last three columns were computed by regressing the baseline characteristics on the treatment variable with union council fixed effects and standard errors clustered at the village level. All variables are defined in Appendix C of the paper.

Table A26: Treatment effects on casual work output

	Levels		Logs	
	Preparer	**Finisher**	**Preparer**	**Finisher**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
Random (R)	-0.650	1.444*	-0.025	0.126***
	(0.796)	(0.875)	(0.025)	(0.045)
Central (C)	-0.662	3.578***	-0.026	0.240***
	(0.880)	(0.799)	(0.027)	(0.041)
Observations	719	719	719	719
Control mean	31.29	21	3.413	2.893
Control SD	7.776	10.65	0.248	0.593
R=C p-values	0.99	<0.01	0.99	<0.01

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable in columns 1-2 is the raw output, while in columns 3-4 it is the log of output. All specifications include control variables selected through post-double selection LASSO (Belloni et al., 2014). Columns do not control for union councils because in many cases, matching between workers happened across union councils. Note that one participant (a Bengali) gets dropped because his monthly savings information is missing but LASSO selects savings as a control.

Table A27: Treatment effects on Log [labor output], by ethnicity

	Ethnic majority		Ethnic minority	
	Preparer	**Finisher**	**Preparer**	**Finisher**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
Random (R)	-0.023	0.121	-0.025	0.137**
	(0.036)	(0.079)	(0.030)	(0.068)
Central (C)	-0.009	0.289***	-0.034	0.213***
	(0.038)	(0.072)	(0.029)	(0.064)
Observations	359	359	360	360
Control mean	3.421	2.879	3.405	2.906
Control SD	0.239	0.608	0.257	0.579
R=C p-values	0.67	<0.01	0.80	0.14

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Columns 1-2 report the log output of ethnic majority workers, and columns 3-4 report the log output of ethnic minority workers. We used the same specification as in Table A26.

Table A28: Treatment effects on raw output, by ethnicity

	Ethnic majority		Ethnic minority	
	Preparer	**Finisher**	**Preparer**	**Finisher**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
Random (R)	-0.571	1.123	-0.651	1.830
	(1.153)	(1.393)	(0.909)	(1.283)
Central (C)	-0.071	4.387***	-1.019	3.159**
	(1.238)	(1.392)	(0.908)	(1.230)
Observations	359	359	360	360
Control mean	31.47	20.93	31.12	21.06
Control SD	7.558	11.22	8.015	10.10
R=C p-values	0.64	<0.01	0.72	0.24

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Columns 1-2 report the raw output of ethnic majority workers, and columns 3-4 report the raw output of ethnic minority workers. We used the same specification as in Table A26.

Table A29: Treatment effects on errors made on the products

		Disaggregated	
	Pooled	**Majority**	**Minority**
Variables	**(1)**	**(2)**	**(3)**
Random (R)	-0.004	-0.007	-0.001
	(0.005)	(0.006)	(0.006)
Central (C)	-0.009**	-0.006	-0.009*
	(0.004)	(0.005)	(0.005)
Observations	719	359	360
=Control mean	0.0420	0.0431	0.0410
Control SD	0.0502	0.0512	0.0494
R=C p-values	0.26	0.85	0.12

Robust SE clustered at village-level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The outcome in all three columns is the proportion of errors made in the casual work task (proportion of errors = number of errors/(number of errors + number of total error-free packets)). Column 1 reports the effect on both Bengalis (majority) and Santals (minority); columns 2-3 report disaggregated effects by ethnicity. We used the same specification as in Table A26.

Table A30: Correlations between log output as finisher and prosociality

	Productivity of Bengalis as a finisher			**Productivity of Santals as a finisher**		
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Altruism (0-100)	0.1040*	-	-	0.1158*	-	-
Solidarity (0-100)	-	-0.0180	-	-	0.0654	-
Trust (0-20)	-	-	0.0773	-	-	0.0601

Note: Pairwise correlations between log productivity as a finisher and prosociality. *** p<0.01, ** p<0.05, * p<0.1.

B Model details

B.1 A simple model: setup and derivations

Here, we present a simple model to help explain some of the features of finishers' productivity displayed in Figure IV, Panel C. The model abstracts away from the role of the preparer and focuses on the finishers. There are two rounds, each with a designated finisher: if a worker is a finisher in the first round, we refer to them as the *first finisher*, while if a worker is the finisher in the second round, we refer to them as the *second finisher*. The second finisher j observes the first finisher i's effort, $j \neq i$, while the reverse is not true. Workers can be one of two ethnicities: $i = B$ (Bengali) and $i = S$ (Santal).

For $i \neq j$, as in our experiment, the team output (number of packets) is given by the sum of efforts of the two finishers:

$$q = e_f^i + e_s^j, \tag{B.1}$$

where e_f^i is the effort of the *first finisher* i and e_s^j is the effort of the *second finisher* j. Note that the subscripts f and s denote the "first" and "second" finishers, respectively. The team receives $2w$ per packet produced, with each member receiving w (equivalent to 2 Taka in our experiment).

The utility function of the first finisher i is defined as:

$$U_f^i = w\left(e_f^i + e_s^j\right) - \frac{1}{2}\left(e_f^i\right)^2 - \frac{1}{2}\theta^i\left(e_f^i - e_s^j\right)^2 + \alpha^i w\left(e_f^i + e_s^j\right), \tag{B.2}$$

while that of the second finisher j is given by:

$$U_s^j = w\left(e_f^i + e_s^j\right) - \frac{1}{2}\left(e_s^j\right)^2 - \frac{1}{2}\theta^j\left(e_s^j - e_f^i\right)^2 + \alpha^j w\left(e_f^i + e_s^j\right), \tag{B.3}$$

where $\alpha^i \geq 0$, and $\theta^i \geq 0$. The first two components of (B.2) and (B.3) represent the individual proceeds from production (i.e., wq) minus the individual cost of effort ($C\left(e_f^i\right) = \frac{1}{2}\left(e_f^i\right)^2$ and $C\left(e_s^j\right) = \frac{1}{2}\left(e_s^j\right)^2$). The third part of the utility function captures workers desire to *conform* to or *match* the effort of the other worker, with θ^i denoting worker i's taste for conformity. Finally, the last component of the utility function captures the *altruistic* concern of a worker toward their partner, where α^i is the intensity or weight a worker i places on the income of worker j.

As stated above, each worker $i = B, S$, can be a first or second finisher. Thus, the game is *sequential* in which the first finisher plays first while the second one plays second.

B.1.1 First case: The Santal worker is the first finisher and the Bengali worker is the second finisher

Using (B.2), the *first finisher Santal*'s utility function is given by:

$$U_f^S = w\left(e_f^S + e_s^B\right) - \frac{1}{2}\left(e_f^S\right)^2 - \frac{1}{2}\theta^S\left(e_f^S - e_s^B\right)^2 + \alpha^S w\left(e_f^S + e_s^B\right), \tag{B.4}$$

while, using (B.3), the *second finisher Bengali*'s utility function is equal to:

$$U_s^B = w\left(e_f^S + e_s^B\right) - \frac{1}{2}\left(e_s^B\right)^2 - \frac{1}{2}\theta^B\left(e_f^S - e_s^B\right)^2 + \alpha^B w\left(e_f^S + e_s^B\right). \tag{B.5}$$

We solve the model using backward induction. Thus, we first solve the *second finisher Bengali*'s problem. The first-order condition is equal to:

$$e_s^B = \frac{\left(1+\alpha^B\right)w + \theta^B e_f^S}{1+\theta^B}. \tag{B.6}$$

Clearly, the effort of the second-finisher Bengali e_s^B is increasing with the first-finisher Santal's expected effort e_f^S, their income w, and their degree of altruism α^B. We can determine e_f^S and show that $\frac{\partial e_f^S}{\partial \alpha^S} > 0$, $\frac{\partial e_f^S}{\partial \alpha^B} > 0$.[1] Similarly, we can calculate e_s^B and show that $\frac{\partial e_s^B}{\partial \alpha^S} > 0$, $\frac{\partial e_s^B}{\partial \alpha^B} > 0$.[2] Finally, since the team output is given by (B.1), that is, $q = e_f^S + e_s^B$, we have $\frac{\partial q}{\partial \alpha^S} > 0$, $\frac{\partial q}{\partial \alpha^B} > 0$. The more the *first finisher Santal* and the *second finisher Bengali* are altruistic and care about each other's income, the higher is their effort and the total quantity produced.

B.1.2 Second case: The Bengali worker is the first finisher and the Santal worker is the second finisher

This case is very similar to the previous case. Relabeling the superscripts we obtain:

$$e_s^S = \frac{\left(1+\theta^S+\theta^B\right)\left(1+\alpha^S\right)w + \theta^S\left(1+2\theta^S\right)\left(1+\alpha^B\right)w}{\left(1+\theta^S\right)^2 + \theta^B}, \tag{B.8}$$

and

$$e_f^B = \frac{\left(1+\theta^S\right)\left(1+2\theta^S\right)\left(1+\alpha^B\right)w + \theta^B\left(1+\alpha^S\right)w}{\left(1+\theta^S\right)^2 + \theta^B}. \tag{B.9}$$

Thus, the team output is now equal to $q = e_f^B + e_s^S$ and we easily obtain $\frac{\partial q}{\partial \alpha^S} > 0$, $\frac{\partial q}{\partial \alpha^B} > 0$.

[1] By plugging (B.6) into (B.4), we obtain

$$U_f^S = w\left(1+\alpha^S\right)\left(\frac{\left(1+\alpha^B\right)w}{1+\theta^B} + \left(\frac{1+2\theta^B}{1+\theta^B}\right)e_f^S\right) - \frac{1}{2}\left(e_f^S\right)^2 - \frac{1}{2}\theta^S\left(\frac{\left(1+\alpha^B\right)w}{1+\theta^B} - \frac{e_f^S}{1+\theta^B}\right)^2.$$

The first-order condition leads to

$$e_f^S = \frac{\left(1+\theta^B\right)\left(1+2\theta^B\right)\left(1+\alpha^S\right)w + \theta^S\left(1+\alpha^B\right)w}{\left(1+\theta^B\right)^2 + \theta^S}. \tag{B.7}$$

[2] By plugging (B.7) into (B.6), we obtain

$$e_s^B = \frac{\left(1+\theta^S+\theta^B\right)\left(1+\alpha^B\right)w + \theta^B\left(1+2\theta^B\right)\left(1+\alpha^S\right)w}{\left(1+\theta^B\right)^2 + \theta^S}.$$

B.2 Interpreting the empirical results in Figure IV

Let us now interpret the results of the field experiment presented in Figure IV Panel C through the lens of our model.

(i) **Figure IV, Panel C:** Treated Bengalis produce more packets than untreated Bengalis, regardless of whether they are first or second finishers. One plausible explanation for this result within our model is that treated Bengalis, whether they are first or second finishers, are more altruistic towards the Santals. This is reflected in a higher weight α^B in their utility function, that is, $\alpha^B_T > \alpha^B_{NT}$, where the subscripts T and NT refer to treated and non-treated, respectively. In our model, when $\alpha^B_T > \alpha^B_{NT}$, the effort exerted by treated Bengali workers is higher compared to non-treated Bengali workers. Note also that the similarity in effort for Bengalis between serving as first or second finishers suggests that θ^B does not play a significant role in their utility function, since they do not seem to exactly match the effort of Santals' when the latter act as first finishers.

(ii) **Figure IV, Panel C**: When Santals are second finishers, their effort is higher when Bengali workers are treated compared to when they are not treated. According to our model, this occurs because when Santals are second finishers, they observe the effort exerted by the Bengali worker before deciding how much effort to exert. Since treated Bengalis put forth more effort as first finisher, second-finisher Santals *match* their effort, resulting in higher team output. However, this is not true when Santals are first finishers because they do not observe the effort of the (second-finisher) Bengali worker. This suggests, according to equation (B.7), that Santals do not anticipate significant variation in the altruism of Bengalis (α^B) across treatments, and therefore their own effort also does not differ significantly across treatments.

C Description and Content Analysis of the Documentary Film

Description. Table C1 summarizes the film's themes and the types of information provided. The three main themes are culture, economic hardship, and the potential for success. Under culture, the topics include traditions and festivals, music and dance, cuisine, and identity and language. Under economic hardship, the topics cover issues related to housing, livelihoods, education, health, gender roles, and various other economic challenges. Finally, the potential for success includes information on educational achievements, professional success, children's aspirations for the future, and the societal contributions of the Santals.

Content Analysis. To examine whether the documentary film differs in the language used for dimensions such as emotive and moral sentiment compared to other media types, such as entertainment media, we analyzed the documentary film's English transcript (i.e., English subtitles) using the LIWC-22 software (Boyd et al., 2022) to quantify such content.[3] LIWC, or Linguistic Inquiry and Word Count, analyzes text to reveal insights about the communication style, psychological state, and social connections expressed through our documentary film.

After uploading the documentary film transcript into LIWC, the software first cleans and prepares the text by removing punctuation, converting to lowercase, and handling formatting issues. Next, it scans the transcript and matches words to categories in its built-in dictionary, which includes dimensions like "positive emotion", "negative emotion", and "moral sentiment". LIWC then calculates the percentage of words that fall into each category, providing a quantitative measure of the prevalence of different themes or psychological constructs. In addition to these categories, the software also analyzes linguistic dimensions like pronoun use and cognitive complexity. Using this output, we created Table C2 that shows the percentage of words in each category, which we then interpret to understand the underlying psychological and social contents highlighted in the documentary.

From Table C2, focusing mostly on the emotional tone, we see that the overall tone of the documentary is slightly negative-to-neutral (with a score of 35.83, on a range from 1-99). However, the use of explicitly emotional language is minimal. In particular, the combined use of sadness-related words, terms expressing negative emotions, or those setting a negative tone is 2.17% of all words used, which suggests that even though the overall tone is slightly negative, the use of negativity and sadness-related words was limited. We also see that the authenticity score is higher (65.93, on a range from 1-99), implying the documentary film shows genuine and personal expression, which is important to understand whether people perceived it as authentic storytelling or fictional. Finally, since we focus on the impacts on prosociality in several dimensions in the paper, we can see that social language constitutes 5.92% of all words used in the documentary, where the use of prosociality-related words is almost negligible (0.65%).

As an additional analysis, we sourced transcripts for other documentary films focusing on Bangladesh. We used the tool called Serper (`https://serper.dev/`) to search for videos relating to the term "Bangladesh objective documentaries YouTube" and focused on the top 100 results available until December 19, 2024. We then used the Python package *YouTube Transcript API* to retrieve transcripts of each of these videos, where 74 Bangladeshi documentary transcripts with English subtitles were successfully extracted. We then uploaded these transcripts along with the transcript of our documentary film to the LIWC software to directly compare the emotional tone and moralization of our documentary film to 74 popular documentary films focusing on Bangladesh on YouTube. In addition, we also compared our documentary's transcript to transcripts of general films that are built in LIWC (which has 19,970 English subtitles, of both English and non-English films, that were released between 1912-2014).

[3]The LIWC-22 website can be accessed via: `https://www.liwc.app/`

Figure C1 shows the comparisons in terms of emotional value. We can see that the emotional value of our documentary film ('Santal Doc' bar) is nearly identical to popular Bangladeshi documentary films ('BGD Doc' bar). Importantly, the emotional content in 'Santal Doc' is much lower compared to the emotional content in general films, suggesting our documentary film is indeed much less sensational than general films. When we focus only on sad content, as shown in Figure C2, we see that sadness content is slightly higher than in 'BGD Doc,' but it is much less than in general films.

Figure C3 compares the use of words reflecting moral judgments across different types of media. For example, this would include instances where a person in the documentary makes a moral evaluation about someone else's behavior or character. We find that such moralization is much less frequent in our documentary compared to other documentaries or general films. This implies that our documentary film presents events and characters without imposing strong moral judgments, taking a more observational approach to storytelling. This allows viewers to form their own opinions about the moral implications of the events depicted.

Table C1: Themes and Subcategories, and Information in the Documentary Film

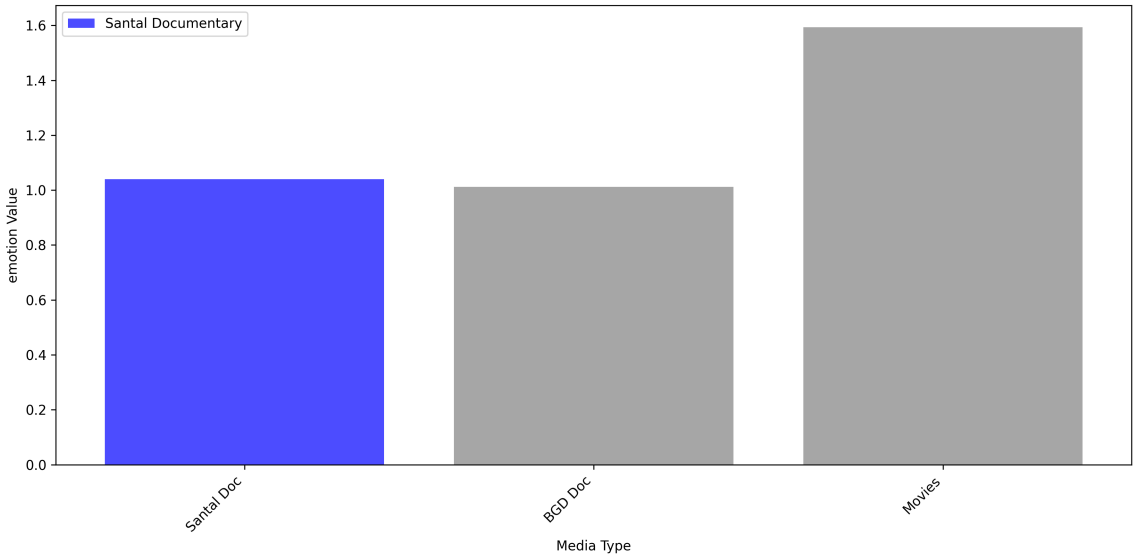
Theme	Subcategory	Information
Culture	Traditions and festivals	The Santals have rich traditions, including festivals like Baha (Blossom Festival) and Shialsei. They celebrate 13 festivals annually and follow ancient customs.
	Music and dance	Traditional songs and dances are integral, passed down through generations. They are used to welcome guests during festivals like Charak Mela.
	Food	Staple foods include rice, fish, crab, pork, and traditional cakes (pithapuli). Meals are communal during festivals.
	Identity and language	They speak Santali at home but face challenges with Bangla in schools and society. Strong emphasis on passing cultural practices through generations.
Struggles and economic hardship	Housing	Many live in dilapidated homes on government land without ownership documents. Vulnerable to storms and monsoon rains.
	Livelihoods	Dependence on agriculture, day labor, bamboo crafting, and cattle raising. Markets are tough, wages are low, and profits are uncertain.
	Education	Children's education is disrupted by poverty, language barriers, and lack of resources like tutors. COVID-19 worsened the situation.
	Health	Limited access to healthcare. Dependence on loans or selling assets for medical treatment. High rate of infant and child mortality.
	Gender roles	Women manage households, fetch water, and assist in income generation through activities like poultry farming.
	Economic challenges	Persistent poverty, reliance on loans, and inability to scale up businesses due to resource constraints. Many businesses shut down due to lack of customers.
Potential for success	Educational achievements	Two students interviewed are studying at public universities. Education is seen as a way out of poverty despite challenges. The importance of providing educational opportunities for the Santals to complete their education.
	Professional success	Interviewed a successful college teacher and an NGO manager who contribute to society and inspire others. Discussions on social barriers and how difficult it was to overcome those.
	Aspirations of children	Children express dreams of finishing education and becoming doctors, nurses, or joining the army. Despite high aspirations, girls find it difficult to attend school due to barriers. Parents try to support these aspirations despite limited means.
	Societal contributions	Santals have contributed during Bangladesh's Liberation War and continue to enrich society through their professions and culture.

Table C2: LIWC Measure Definitions and Findings

Measures	**Definitions**	**Scores from LIWC**
Word count	Total number of words analyzed, which is the length of our film transcript	3,734 words were analyzed
Analytic	Reflects analytical or logical thinking, where a low score indicates a more narrative or intuitive style	24.32
Clout	Measures confidence or authority in language. Higher scores indicate confident and influential communication	72.95
Authentic	Measures perceived honesty and personal authenticity. Higher scores indicate genuine and personal expression	65.93
Tone	Reflects emotional tone, where a lower score suggests negativity or anxiety, while a higher score indicates positivity or optimism	35.83
Drive: Affiliation	Social connection-related words	7.61%
Drive: Achieve	Achievement-related language	2.52%
Drive: Power	Words indicating dominance or control	0.99%
Affect: Tone Positive	Emotional language: positive tone	2.73%
Affect: Tone Negative	Emotional language: negative tone	1.63%
Affect: Emo Positive	Emotional language: positive emotion	0.46%
Affect: Emo Negative	Emotional language: negative emotion	0.51%
Affect: Emo Sadness	Emotional language: sadness related words	0.03%
Cognition	Reflects cognitive engagement	13.87%
Social	Broad social language	General (2.87%), Prosocial (0.64%), Family (2.3%), Friend (0.11%)
Culture	Cultural or political references	0.54%
Lifestyle	References to daily life topics like home, work, or leisure	7.2%
Physical	Health or body-related terms	2.73%
Food	References to food or eating	1.34%
Time	General time-related language	5.3%

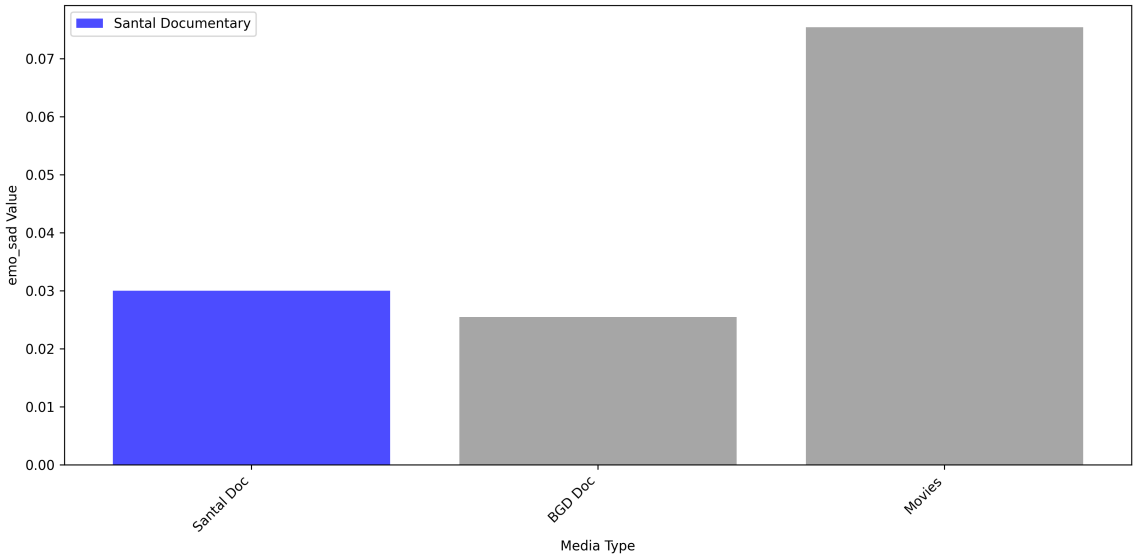
Notes: Scores without % are in the range from 1 to 99. We excluded categories such as the use of pronouns and punctuations, words per sentence, use of big words, prepositions, verbs, adverbs, auxiliary verbs, conjunctions, adjectives, use of numbers, etc., as they are not useful to understand the emotional and moral contents of the documentary film.

Figure C1: Emotion value comparisons, using LIWC



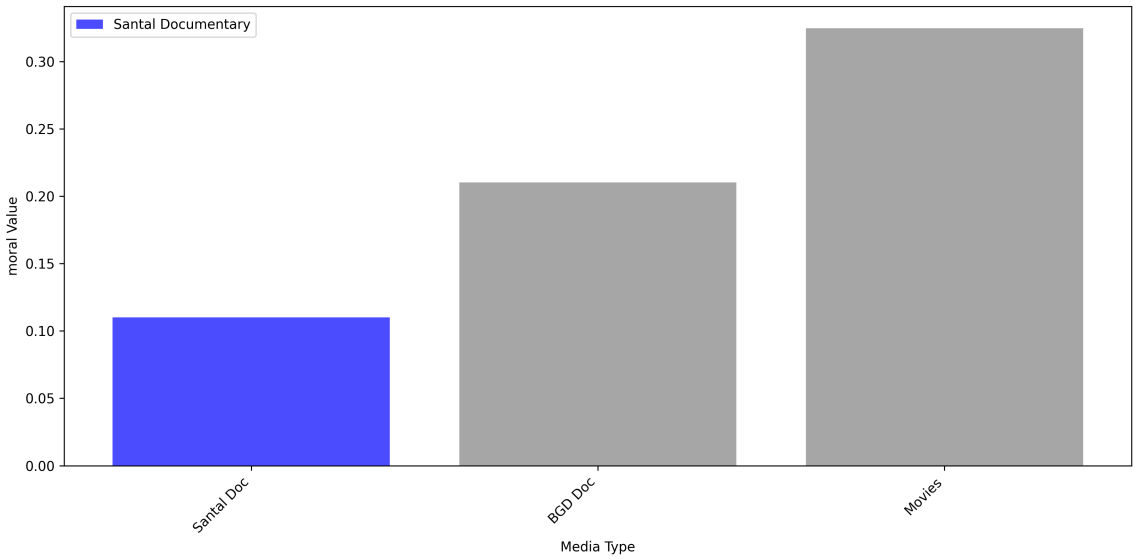
Note: The overall 'emotion' category, a subcategory of Affect, contains both positive and negative words. We compare the mean emotion value of the Santal documentary film ('Santal Doc') computed using LIWC to the mean values for (i) 'BGD Doc', which includes other Bangladesh documentaries taken from YouTube, and, (ii) 'Movies', which is from general films. A higher value/bar corresponds to higher emotional content.

Figure C2: Sadness comparisons, using LIWC



Note: The overall 'sadness' category, a subcategory of Affect, contains words relating to sadness such as 'sad, disappoint*, cry*', etc. We compare the mean sadness value of the Santal documentary film ('Santal Doc') computed using LIWC to the mean values for (i) 'BGD Doc', which includes other Bangladesh documentaries taken from YouTube, and, (ii) 'Movies', which is from general films. A higher value/bar corresponds to higher sad content.

Figure C3: Moral value comparisons using LIWC



Note: The overall 'moral' category contains words reflecting moral judgments about another behavior such as 'wrong, honor*, deserv*, judge', etc. We compare the mean moral value of the Santal documentary film ('Santal Doc') computed using LIWC to the mean values for (i) 'BGD Doc', which includes other Bangladesh documentaries taken from YouTube, and, (ii) 'Movies', which is from general films. A higher value/bar corresponds to higher moral content.

D Variable Descriptions and Experimental Instructions

D.1 Variable descriptions

Altruism. We use a standard dictator game to measure altruism towards non-coethnics (Eckel and Grossman, 1996). Each participant (Player A) was paired with a non-coethnic recipient from their own village (Player B). This arrangement paired a majority member with a minority member, and vice versa. Player A received an endowment of 100 Taka (=$1) and was asked to privately divide it between themselves and Player B. The share transferred to Player B (through an envelope) measures Player A's altruism. This game was played individually with an enumerator at Player A's home, and all players remained anonymous. Out of the three measures of prosociality, this was always measured first.

Solidarity. We used a simplified solidarity game to measure solidarity and expected solidarity towards non-coethnics during unexpected shocks (Selten and Ockenfels, 1998). This game was played at the decision maker's home (while the recipient stayed at their own home), and both players remained anonymous. Again, this game paired a majority member with a minority member, and vice versa. Each player was given an endowment of 100 Taka prior to the game. The enumerator tossed a coin in front of Player A; if it landed on 'heads', Player B's entire endowment was destroyed, while if it landed on 'tails', Player B's endowment remained intact. Before the coin toss, Player A decided how much to give to Player B in the event of a 'heads' result. The money was placed in an envelope in private and handed to the enumerator, who then tossed the coin. If 'heads', the envelope was kept by the enumerator to later give to Player B. If 'tails', the envelope was returned to Player A. Therefore, the amount offered to Player B measured Player A's solidarity towards non-coethnics who lost their wealth due to a negative shock. Before the coin toss, Player A was asked, "If you play this game with Player B, how much do you think they will be willing to give to you?" This question measured their level of expected solidarity; however, note that this was self-reported, and enumerators recorded their responses. Out of the three measures of prosociality, this was always measured second.

Trust. At the first endline, we were also interested in the formation of interethnic trust following the intervention, measured using a survey question (Falk et al., 2018). We asked "I assume that Santals/Bengalis have only the best intentions", which was answered on a scale 0-20, where 0 means 'not agree at all' and 20 means 'completely agree'. Santal respondents were asked about trust towards Bengalis, and Bengali respondents were asked about trust towards Santals. A thermometer scale was used, with participants indicating their response by pointing using their finger and enumerators recording the value.[1] We introduced this scale to improve precision. Out of the three measures of prosociality, this was always measured last.

Stereotypes index. Based on the following 6 questions that capture Bengalis' stereotypes about Santals (each answered on a 0-10 scale, where 10 means "completely agree", thus the aggregated score is between 0 and 60):

1. Santals are often unclean/unhygienic
2. Santals would make very good doctors *(reverse scoring required)*
3. I have not met or known any Santals who have established themselves or made a mark
4. Santals do not make very good school teachers.
5. Santals do not continue beyond schools

[1] Note that baseline trust was measured verbally (without a thermometer scale) on a 0-10 scale.

6. Santals should continue working in the agricultural sector

Discriminatory opinions index. Based on the following 7 questions that capture Bengalis' opinions about Santals (each answered on a 0-10 scale, where 10 means "completely agree", thus the aggregated score is between 0 and 70):

1. All Santals that I know are honest people *(reverse scoring required)*
2. I always eat food and drinks offered by Santals *(reverse scoring required)*
3. Schools should be separate for Bengali and Santal children
4. I think there should be more Santal teachers in my child's school *(reverse scoring required)*
5. I do not enjoy working/doing business with Santals
6. Some of my children's best friends are Santals *(reverse scoring required)*
7. One can easily trust a Santal person *(reverse scoring required)*

Number of visits to Santal/Bengali neighbors (two variables). Participants were asked "How many times do you visit your Santal/Bengali neighbors in a month?". We directly use this frequency as the outcome.

Number of Santal/Bengali visitors (two variables). Participants were asked "How often do your Santal/Bengali neighbors visit you in a month?". We directly use this frequency as the outcome.

Social connections with network-central Bengalis (two variables). Because we have both network central (7 per village) and randomly selected ethnic majority participants (7 per village) in the *Central* arm, we collected data on how connected each Bengali untreated respondent (7 per village) is to each network central in this treatment arm. Specifically, we asked: "How often did you visit .....'s house or s/he visited your house last month?", where ..... is the name of network central participant. Then we asked: "How long does it take you (in minutes) to walk to .....'s house?".[2] We ask these questions 7 times to each Bengali untreated respondent, one for each network central participant in the *Central* arm. Using these responses, we created two variables for each Bengali untreated respondent: (1) number of visits to network-central, and (2) average distance to network-central.

Interaction index (two variables—one interactions with Bengalis and another with Santals). Based on the following 4 questions that capture how often Bengalis (Santals) interact with their Santal (Bengali) neighbors (answered as *yes=1* or *no=0*, thus the aggregated score is between 0 and 4):

1. Do you offer them food when they visit you?
2. Do you offer them chair/seat when they visit you?
3. Do you invite them during festivals?
4. Do they invite you during festivals?

Intercultural competence index. The simplified version of this index (4-items), used in Siddique and Vlassopoulos (2020), captures a Bengali's awareness of Santals and their culture. The four questions are as follows:

1. What is the language spoken by Santals?

[2] We wanted to record GPS coordinates of each respondent during the survey but we could not do it because of poor 3G internet coverage in these villages.

2. Do you speak or understand that language?

3. What is their major religion?

4. Name a major Santal festival?

Answering each question correctly gave 1 point, with a maximum total of 4. Therefore, the score is between 0 and 4, where 4 means 'full competence'. This question was only asked to Bengalis.

Number of ethnic minority close friends. Participants were asked "Name your ten closest friends (full name)". Based on surnames, enumerators counted the number of ethnic minority close friends they have and recorded that value.

Water bill charge to non-coethnics. Participants were asked "We know you need to pay water bills. How much do Santals/Bengali pay per fetch when they come to fetch water at your house?". We directly use this monetary amount as the outcome. Note that this question was only asked to respondents who had a tubewell installed at their home.

Mental health index. We measured depressive and general anxiety symptoms using the PHQ-4 questionnaire (Kroenke et al., 2009). We asked the following questions: "Over the last two weeks, how often have you been bothered by the following problems? *Not at all=0 / several days=1 / more than half the days=2 / nearly everyday=3*":

1. Feeling nervous, anxious, or on edge.

2. Not being able to stop or control worrying.

3. Feeling down, depressed, or hopeless.

4. Little interest or pleasure in doing things.

We created an index using the standardizing procedure explained before.

Subjective well-being index. We measure this outcome using the following 4 questions from the World Values Survey: "On a scale from 0 to 10, where 0 means *not at all* and 10 means *extremely or all the time*":

1. Overall, how satisfied are you with your life nowadays?

2. Overall, to what extent do you feel the things you do in your life are worthwhile?

3. Overall, how happy do you feel nowadays?

4. Overall, how anxious do you feel nowadays? *(reverse scoring required)*

We created an index using the standardizing procedure explained before.

Food security index. We used the USDA Household Food Security Survey Module to measure food security in the household of Santals only, each question answered on a 3-point scale: *often true=2, sometimes true=1, never true=0*; thus, the score is between 0 and 6. We ask the following questions "In the last 6 months, can you tell me if these statements were true for you?:

1. We worried whether our food would run out before we got money to buy more.

2. The food that we bought just didn't last, and we didn't have money to get more.

3. We couldn't afford to eat balanced meals.

We created an index using the standardizing procedure explained before.

New employment. Did you start any new job recently that lets you earn more than before? Coded as *yes=1* and *no=0*. We asked this question to Santals only at endline.

Monthly income. Last months' household income. We directly use this monetary amount and $ln(income)$ as the outcome. We asked this question to Santals only at endline.

Prosociality at baseline. Following Falk et al. (2018), we measured altruism and trust using the following questions. We also measured solidarity in a similar fashion. We asked:

1. *Altruism:* Imagine the following situation—Today you unexpectedly received 5,000 Taka. How much of this amount would you donate to a Santal/Bengali person in your village? We directly used this monetary amount as the outcome.

2. *Solidarity:* Imagine your neighbor, Hopna Kisku/Iqbal Rahman, lost his house during a heavy storm. Also, imagine that today you unexpectedly received 5,000 Taka. How much of this amount would you donate to Iqbal Rahman and his family? We directly used this monetary amount as the outcome.

3. *Trust:* Please indicate your answer on a scale from 0 to 10. A 0 means *does not describe me at all*, and a 10 means *describes me perfectly*. "I assume that Santal/Bengalis have only the best intentions." We directly used this score as the outcome.

Interpersonal Reactivity Index. To measure individual empathy, we use the Interpersonal Reactivity Index by Davis (1983). We measure three dimensions of empathy to understand which channels can get activated from the documentary film: perspective-taking or PT (the tendency to spontaneously adopt the psychological view of others in everyday life), empathic concern or EC (the tendency to experience feelings of sympathy or compassion for unfortunate others), and personal distress or PD (the tendency to experience distress or discomfort in response to extreme distress in others). We ask the following:

"The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing between 0 and 4, where 0 means *does not describe me at all* and 4 means *describes very well*. Thus, the higher the number, the more it describes your thoughts and feelings. Answer as honestly as you can.":

1. I often have tender, concerned feelings for people less fortunate than me. (EC)
2. I sometimes find it difficult to see things from the "other guy's" point of view. (PT) (-)
3. Sometimes I don't feel very sorry for other people when they are having problems. (EC) (-)
4. In emergency situations, I feel apprehensive and ill-at-ease. (PD)
5. I try to look at everybody's side of a disagreement before I make a decision. (PT)
6. When I see someone being taken advantage of, I feel kind of protective towards them. (EC)
7. I sometimes feel helpless when I am in the middle of a very emotional situation. (PD)
8. I sometimes try to understand my friends better by imagining how things look from their perspective. (PT)
9. When I see someone get hurt, I tend to remain calm. (PD) (-)
10. Other people's misfortunes do not usually disturb me a great deal. (EC) (-)

11. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (PT) (-)
12. Being in a tense emotional situation scares me. (PD)
13. When I see someone being treated unfairly, I sometimes don't feel very much pity for them. (EC) (-)
14. I am usually pretty effective in dealing with emergencies. (PD) (-)
15. I am often quite touched by things that I see happen. (EC)
16. I believe that there are two sides to every question and try to look at them both. (PT)
17. I would describe myself as a pretty soft-hearted person. (EC)
18. I tend to lose control during emergencies. (PD)
19. When I'm upset at someone, I usually try to "put myself in his shoes" for a while. (PT)
20. When I see someone who badly needs help in an emergency, I go to pieces. (PD)
21. Before criticizing somebody, I try to imagine how I would feel if I were in their place. (PT)

Statements with (-) require reverse scoring. There are 7 questions on each of the three dimensions, so the score for each would be between 0 and 28. Thus, a higher aggregate score corresponds to someone being more empathetic.

Social desirability bias index. To address social desirability bias (SDB) concerns pertaining to self-reported outcomes, we closely follow Dhar et al. (2022). Using this 13-point SDB scale (each item is answered as agree/disagree, where socially desirable answers are coded as 1 and 0 otherwise), we aggregate all 13 responses such that the value range is between 0 and 13, where 13 means the respondent is most likely to give socially desirable responses. We asked the following questions at baseline to all participants: "Please answer as accurately as possible if the following characteristics describe you or not. Please answer as either *agree* or *disagree*":

1. It is sometimes hard for me to go on with my work if I am not encouraged (Disagree)
2. I sometimes feel resentful when I don't get my way (Disagree)
3. On a few occasions, I have given up doing something because I thought too little of my ability (Disagree)
4. There have been times when I felt like rebelling against people in authority even though I knew they were right (Disagree)
5. No matter who I'm talking to, I'm always a good listener (Agree)
6. There have been occasions when I took advantage of someone (Disagree)
7. I'm always willing to admit it when I make a mistake (Agree)
8. I sometimes try to get even rather than forgive and forget (Disagree)
9. I am always courteous, even to people who are disagreeable (Agree)
10. I have never been irked when people expressed ideas very different from my own (Agree)

11. There have times when I was quite jealous of the good fortune of others (Disagree)

12. I am sometimes irritated by people who ask favors of me (Disagree)

13. I have deliberately said something that hurt someone's feelings (Disagree)

The socially desirable responses are given in parentheses. Therefore, for each question, a respondent gets a point if their response matches with the socially desirable response and 0 otherwise.

D.2 Experimental instructions

We will play two games involving real money. After playing both games, we will do a coin toss: 'head' will mean you will be paid according to the outcome of the first game, and 'tail' will mean you will be paid according to the outcome of the second game. The money you win today will be yours to keep.

First game—Dictator: You are playing this game with a Santal/Bengali that lives in your village. For simplicity, let's call this person your opponent. Note that, I will not reveal her/his identity to you, and nor will I reveal your identity to her/him. So, your identity will remain a secret and will never be revealed.

For this game, I will give you 100 Taka in nineteen 5 Taka coins and five 1 Taka coins. I have not given any money to your opponent. You now have to decide how much of the 100 Taka you would like to share with your opponent. You may share either some, all, or nothing with your opponent. Whatever amount you share will be sent to your opponent, and you can keep the remaining amount.

Here is an envelope. Whatever amount you wish to share should be left in the envelope. I will now turn my back.

Once our games end, I will give your envelope to your opponent.

Second game—Solidarity: This game depends on luck. You are playing this game with a Santal/Bengali that lives in your village. For simplicity, let's call this person your opponent. Note that, I will not reveal her/his identity to you, and nor will I reveal your identity to her/him. So, your identity will remain a secret and will never be revealed.

For this game, I will give you 100 Taka in nineteen 5 Taka coins and five 1 Taka coins. I will then do a coin toss. If the coin shows a 'head', then your opponent will lose the entire 100 Taka I gave her/him. That is, in case of a 'head', s/he will have to return the 100 Taka I gave her/him. However, if it is a tail, then s/he can keep the 100 Taka. Is this clear?

Now before I do the coin toss, you have to decide how much of your 100 Taka you would like to give to your opponent if s/he loses her/his money when there is a 'head'. You may pledge to give either some, all, or nothing to your opponent. Remember, your money will only be handed over to your opponent if the coin toss is a 'head'. If it is a 'tail', this money that you pledge would be returned to you. In case of a 'head', whatever amount you pledge will be sent to your opponent, and you can keep the remaining amount.

Here is an envelope. The amount you pledge to give in case of a 'head' should be left in the envelope. I will now turn my back.

If I play this game with your Santal/Bengali opponent, how much do you think your opponent would pledge to give you? (Record this value)

(After s/he hands back the envelope, do the coin toss) Once our games end, I will give your envelope to your opponent.

(After coin toss, do the trust measure.)

D.3 Data collection logistics

We preregistered that, at endline, the lab-in-the-field experiments would be integrated with the survey, with survey-based outcome measures and incentivized experiments presented in two separate blocks, and their order randomized. In implementation, while the order was indeed randomized across respondents, the survey team was unable to program this into the survey software. As a result, the two lab-in-the-field games and the trust question were always conducted and recorded using pen and paper. For this, enumerators were given a matrix where they recorded these decisions. Note also that this randomization occurred at the village level rather than at the individual level: for example, in some villages, participants first completed the survey and then participated in the games and trust measures, while in others, the reverse order was followed. However, we did not systematically record the order in which each village completed the two blocks.

In each village, the survey component was carried out by a team of four enumerators. Two enumerators were assigned to the fourteen main Bengali respondents, one to the untreated Bengali sample, and one to the untreated Santal sample—each enumerator thus interviewed, on average, seven participants per village. In addition, two dedicated 'experimenter' enumerators were responsible for conducting the two lab-in-the-field games and trust measurement across all villages. Since both games and the trust question were quicker than the full survey, a single experimenter typically conducted both games and the trust question in each village. However, s/he typically spent the entire day in a single village.

Each enumerator was provided with a detailed list for each assigned village, which included the village name and ID, the name and ID of the union it belonged to, treatment or control assignment, village size, and ethnic composition (based on information provided by *Ashrai*). Village and union IDs were generated based on alphabetical order. Some village IDs appear missing from the final dataset, as they are not numbered from 1 to 121 continuously. This is because, in a few cases, data collection could not proceed due to local police advising the team to avoid certain villages on the day of the visit due to tensions or security concerns. Enumerators simply skipped those villages and proceeded to the next ones on the list.

During data collection, enumerators recorded information mentioned in their assignment sheets. They were also given unique respondent IDs in advance. For example, in a given village, Enumerator 1 would always use the first seven IDs for the main Bengali sample, Enumerator 2 the last seven IDs of the main sample, Enumerator 3 the seven IDs for untreated Bengalis, and Enumerator 4 the seven IDs for untreated Santals. Participant IDs were constructed systematically using the village ID, respondent group (main Bengali, untreated Bengali, or untreated Santal), and a unique number from 1 to 28 within each village. While each enumerator was assigned seven respondents per village, this was not always the case due to variation in attrition. It was common for enumerators with fewer respondents to assist others to ensure full coverage.

In the casual work task involving paper bag preparation, about ten majority–minority pairs participated each day: five pairs in the morning shift and five in the afternoon shift. Two 'experimenter' enumerators (the same who conducted the games and trust measures at endline) were responsible for overseeing the task alongside the paper bag supplier and for recording productivity measures. All measurements were recorded using pen and paper. For this, enumerators were given a matrix where they recorded the productivity.

E Details on the qualitative evidence

We included additional qualitative questions immediately after Bengalis watched the documentary film to investigate their intentions to help the Santals and their planned actions. We asked open-ended questions focused on Bengalis' willingness to support Santals, the specific forms of support they would consider, their reasoning, and their intentions to encourage other Bengalis in their villages. We specifically asked, *"Now imagine a Santal in your village is in similar conditions as the Santals in the film that you just watched...(1) What would you do in that situation? (2) Why? (3) What would you advise your neighbors to do in that situation?"* Since these questions were only asked to the viewers who watched the documentary on Santals, we cannot claim the following qualitative evidence to be causal.

We received 31 unique responses regarding intentions to help Santals, 18 about underlying reasons for those intentions, and 17 on advice Bengalis would give their neighbors regarding supporting or helping Santals. It is important to note that while respondents may have articulated a wider range of thoughts, enumerators were instructed to record only a limited number of responses or keywords to reduce the data collection burden. Therefore, even though the actual answers could have been more diverse and unique, based on the enumerators' judgment and the recording constraints, we found only these many unique responses.

To further narrow down the responses, we used an Application Programming Interface (API) to access the GPT-4 model (Large Language Models or LLM). This allowed us to programmatically assign the responses into broader categories. This approach aims to use the tool as a replacement for human labelling where the final output is a dataset that can be critically examined and falsified by a human or another LLM. To make this categorization both interpretable and falsifiable we ask for—(i) a suggested category and the keywords and phrases used to make the categorization; (ii) what the LLM thinks the definition of these keywords are in the context of the answers; and, (iii) a more detailed reasoning for the categorization explaining in detail how it reached this answer. To get the response categories, we ran our prompts five times. See Appendix F for more details on the categorization and the prompts used.

Based on this exercise, we identified five broad categories of information using the three questions: (1) General support, encompassing non-specific intentions to help Santals (e.g., "I want to help them"); (2) Economic and financial assistance, focused on intentions to provide financial support or help Santals secure better jobs (e.g., "I want to help them financially"); (3) Advisory support, covering intentions to offer Santals practical or economic advice (e.g., "Give them advice"); (4) Humanitarian reasons (e.g., "Because they are suffering"); and finally, (5) Missing, which includes responses that were empty, irrelevant, or indicated an unwillingness to answer.

We visualize the results on intentions to help Santals using a heatmap in Figure E1. General support intentions were most common (about 55% of participants), followed by economic and financial help (about 21%). Advisory support accounted for only 4%, and approximately 20% of responses were missing. These intention patterns were very similar across the 'Random' and 'Central' treatment arms. When exploring the reasons behind these intentions (see Figure E2), we find that approximately 58% of participants cited humanitarian reasons (e.g., "because they are suffering," "because they are in poverty"). 21% mentioned they needed economic or financial help. The remaining 21% did not provide any response. Finally, when asked what they would advise or encourage their neighbors to do if the Santals in their villages were suffering (see Figure E3), about half of the participants said they would advise them to help Santals (i.e., general support intentions). Economic and financial help (30%) and offering better advice to Santals (4%) were the next most common responses. Again, these responses are very similar across 'Random' and 'Central' arms.

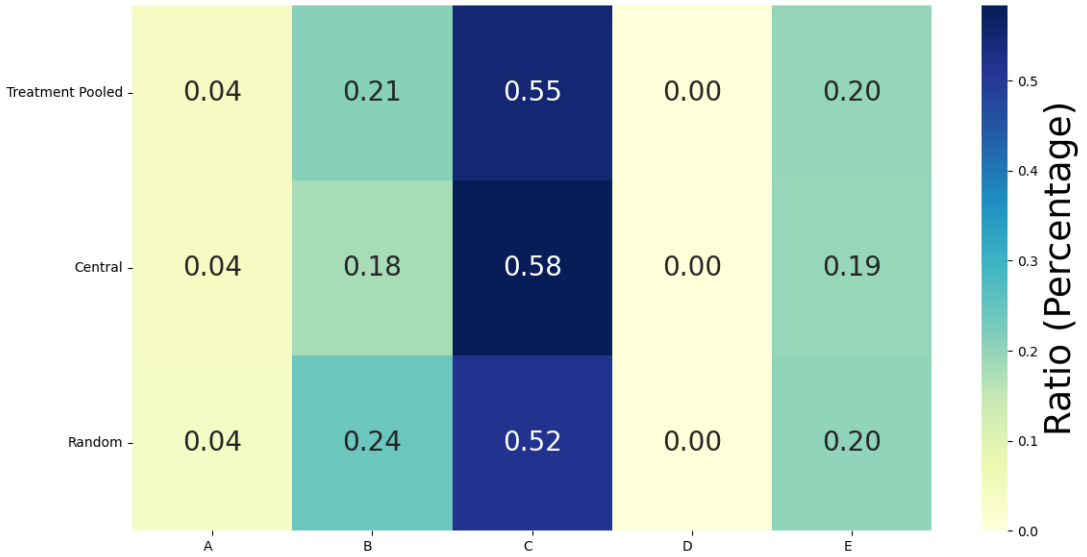
Using Sankey diagrams, we also map the new information received from the documentary

film (discussed in Section 4.1) into the Bengali peoples' intentions to help Santals. In all diagrams, it is clear that almost all newly received information played some role in developing intentions and motivation among Bengalis to help Santals (Figure E4), the reasons behind developing such intentions (Figure E5), and their positive intentions to also encourage their neighbors to help Santals (Figure E6).

We corroborate the above qualitative evidence (which is not causal) with the following quantitative question, which was asked to all participants (including Bengalis in the 'Control' arm): *"How much do you want to help others"* (without specifying the Santals)? Because we can estimate the treatment effect, we interpret this result as causal. We report this estimate in Column 6 in Table I, which shows that Bengalis in both 'RR' and 'CC' arms are significantly more willing to help others.

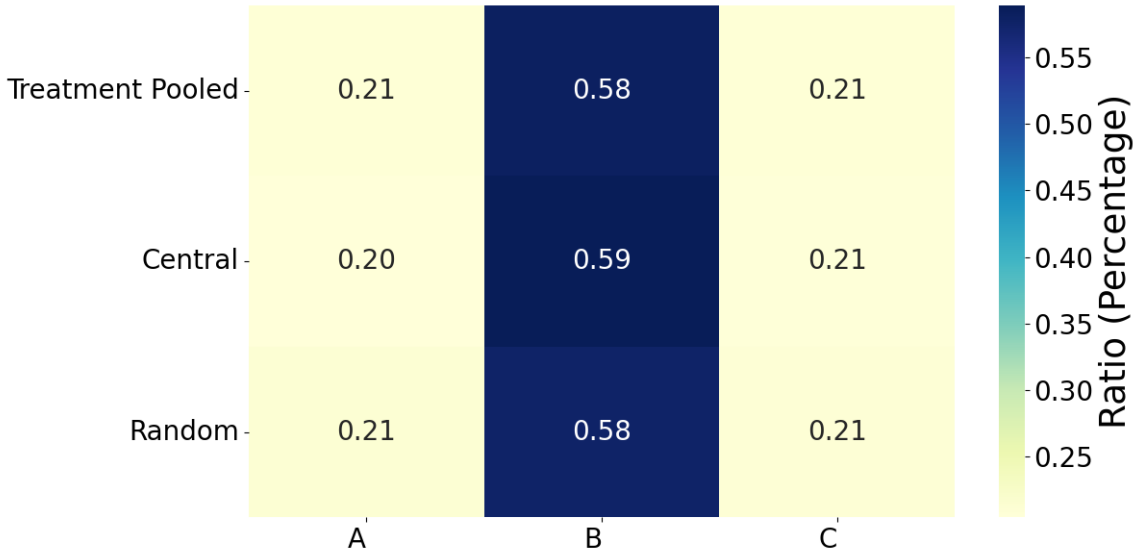
Overall, this analysis highlights the documentary film's potential to promote positive change among viewers.

Figure E1: Heatmap showing the distribution of intentions to help



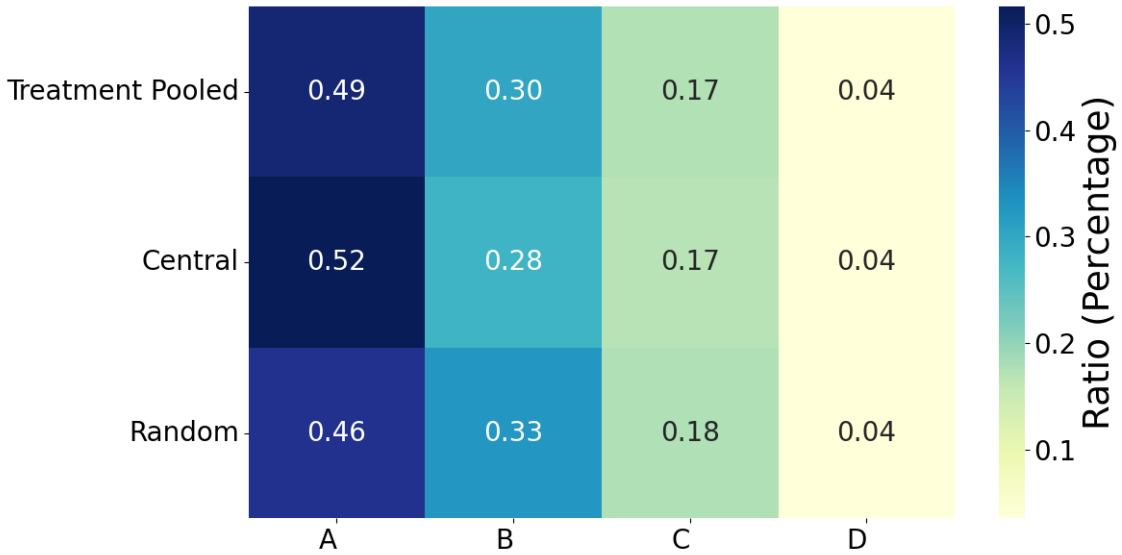
Note: Categories are the top weighted category for distinct comments aggregating results from five iterations of the GPT-4 model. Weight is the count of the number of times GPT-4 predicts a category for each comment divided by five. In all cases, there is only a single category with a weight above 0.5. There are no cases where a category is below 0.5. Categories not included in the heatmap have no comments where they are the top predicted category. Category to Alphabetic Label Mapping: Advisory Support is A, Economic and Financial Aid is B, General Support Intentions is C, Humanitarian Reasons is D, Missing is E.

Figure E2: Heatmap showing the distribution of reasons behind the intentions to help



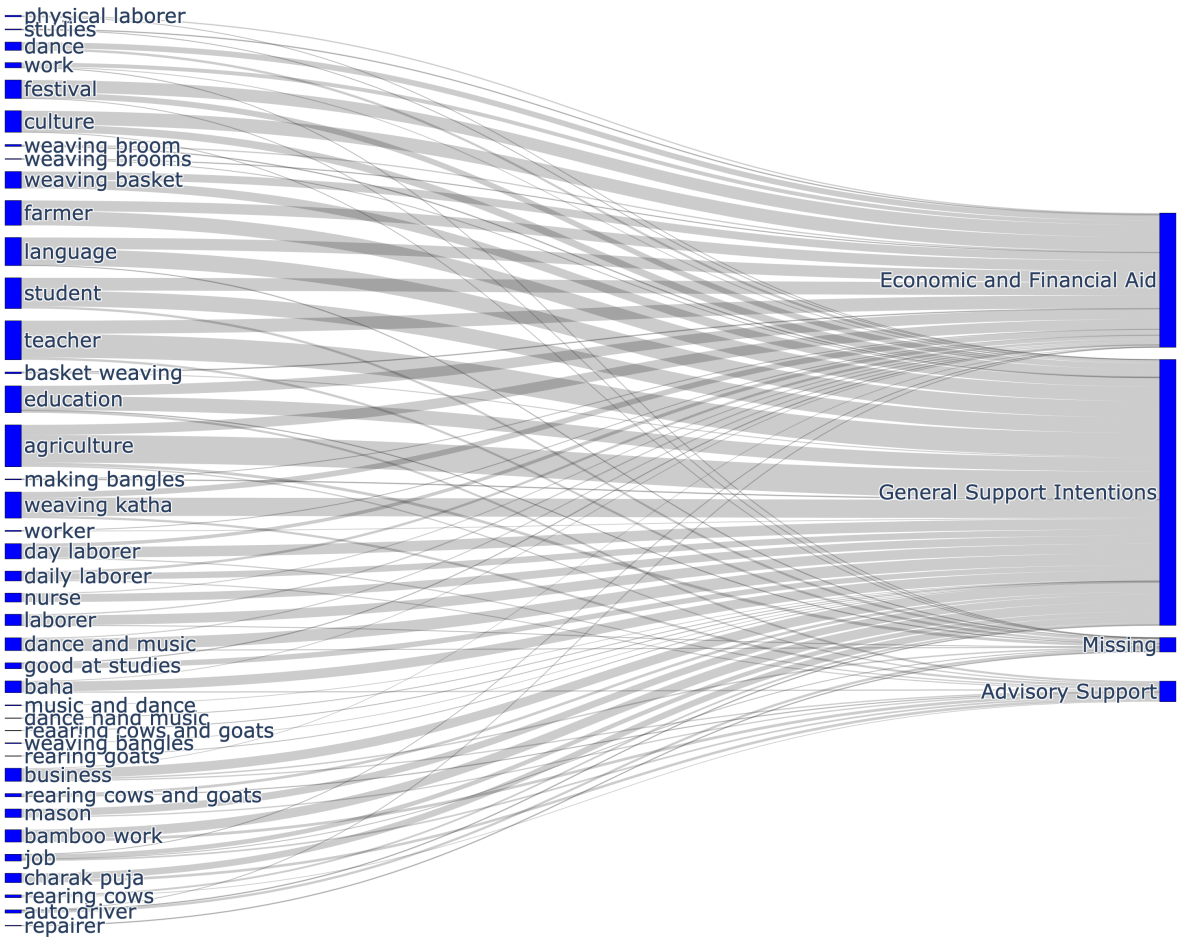
Note: Categories are the top-weighted category for distinct comments aggregating results from five iterations of the GPT-4 model. Weight is the count of the number of times GPT-4 predicts a category for each comment divided by five. In one case, there are two categories with a weight above 0.5. The comment 'should help them because they are living in poverty' has the suggested categories 'Humanitarian Reasons' and 'General Support Intentions' in four out of five iterations for both categories. There are no cases where a category is below 0.5. Categories not included in the heatmap have no comments where they are the top predicted. Category to Alphabetic Label Mapping: Economic and Financial Aid is A, Humanitarian Reasons is B, Missing is C.

Figure E3: Heatmap showing the distribution of the intentions to encourage neighbors when Santals are suffering



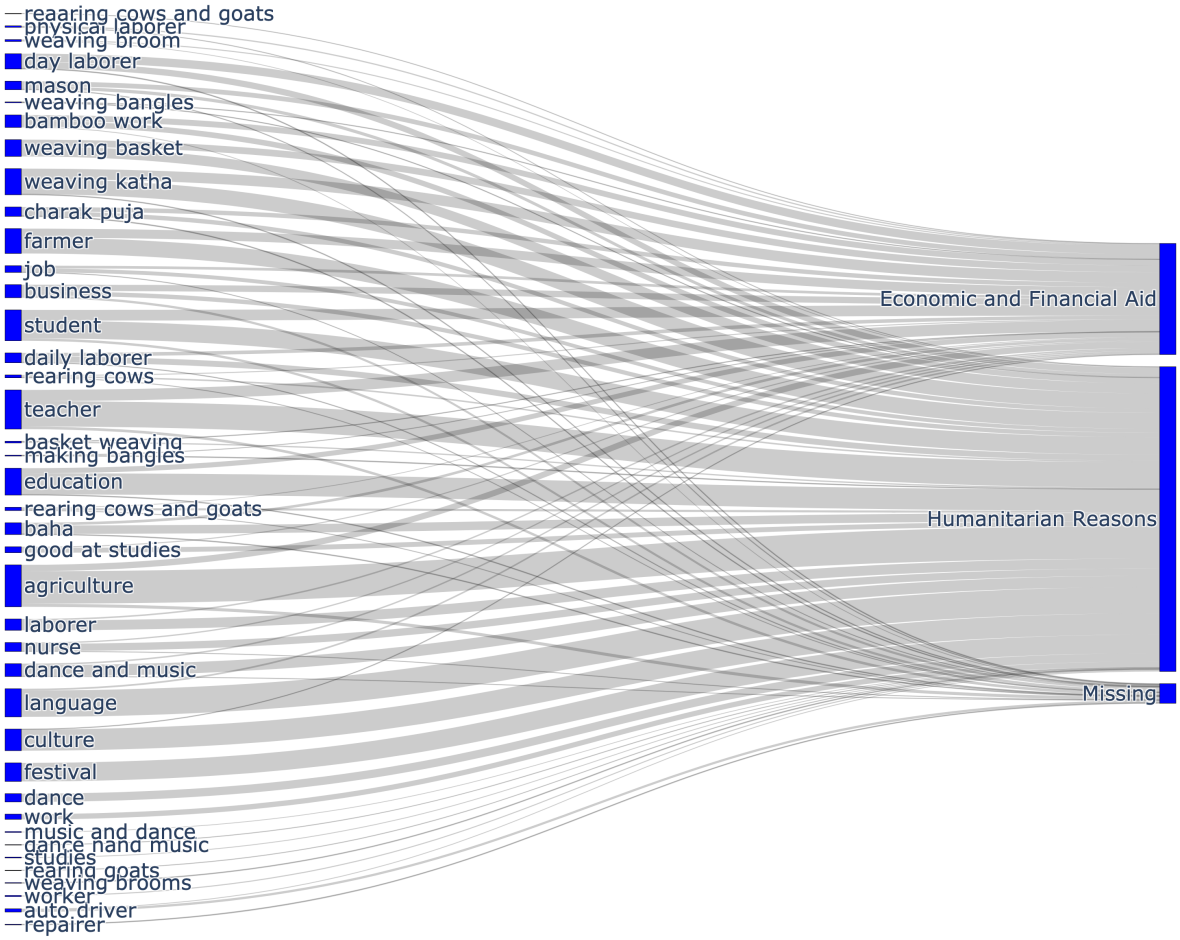
Note: 50 percent of answers fall in the category of general support intentions. Category to Alphabetic Label Mapping: General Support Intentions is A, Economic and Financial Aid is B, Missing is C, Advisory Support is D.

Figure E4: Mapping new information into intentions to help Santals



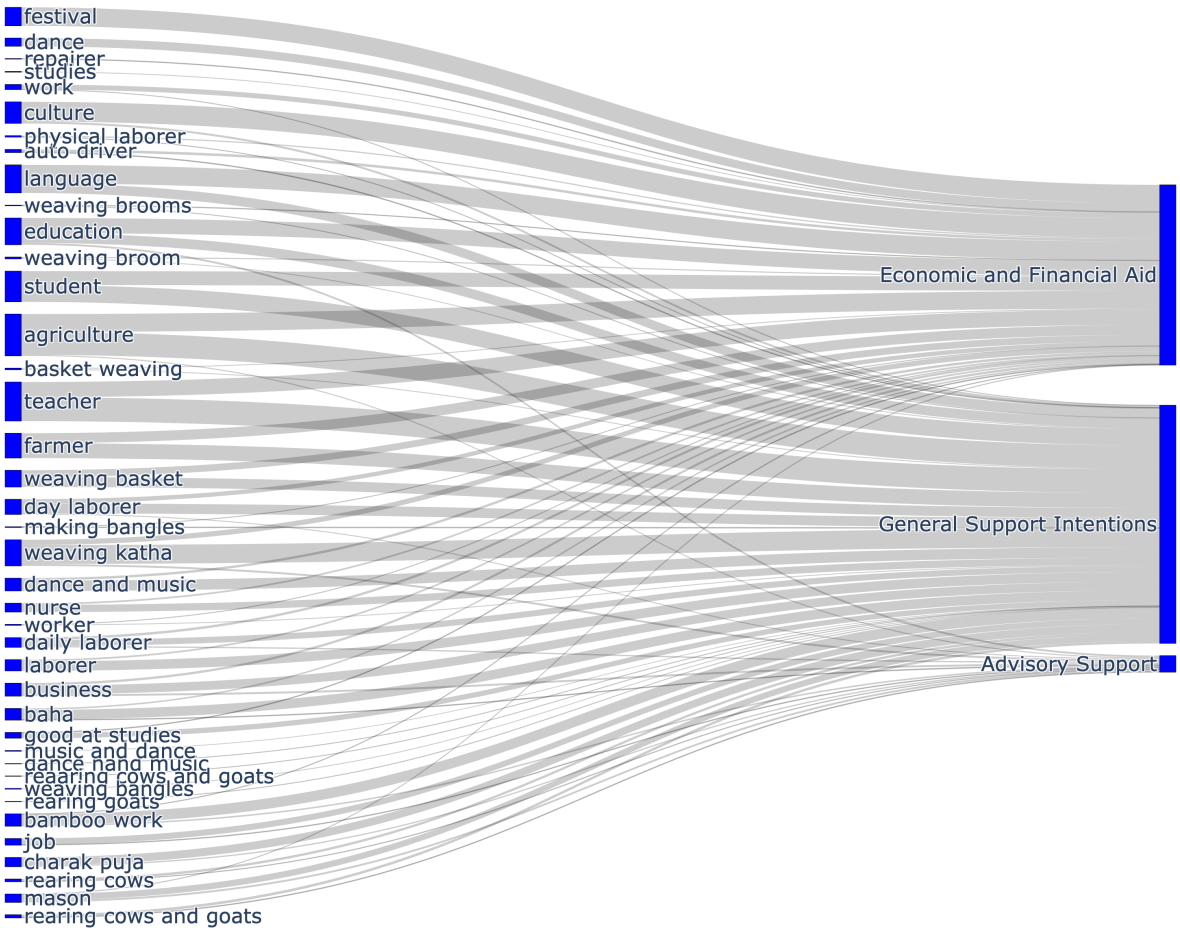
Note: Sankey diagram showing the mapping of new information received through the documentary film (using all unique answers) to Bengalis' intentions to help Santals.

Figure E5: Mapping new information into reasons behind the intentions to help Santals



Note: Sankey diagram showing the mapping of new information received through the documentary film (using all unique answers) to Bengalis' reasons behind the intentions to help Santals.

Figure E6: Mapping new information into the intentions to encourage neighbors when Santals are suffering



Note: Sankey diagram showing the mapping of new information received through the documentary film (using all unique answers) to Bengalis' intentions to encourage neighbors when Santals are suffering.

F Categorization using GPT-4

We use GPT-4 to perform the categorization in a process similar to human labeling. GPT-4 is a distinct service from ChatGPT. GPT-4 allows for programmatic queries over a dataset using common programming languages such as Python. This is performed using an Application Programmatic Interface (API). An API is a method for computers to programmatically access a service. OpenAI makes the 'model weights' available to query through this process. The process involves sending a 'prompt' with categories and the survey response you want to categorize to their model through an API, the model predicts the category for each response and returns it to your computer. To ensure the answers are consistent, we run the query five times and take the most frequently predicted category for each survey response.

A standard approach in survey analysis involves using third-party human labelers for data categorization and analysis (Bochkay et al., 2023). ChatGPT and GPT-4 (the model behind it) have emerged as a viable alternative to human labeling. It significantly outperforms mechanical Turk workers across a variety of labeling tasks and is up to 30 times more cost-effective than using human labelers (Gilardi et al., 2023). Moreover, Veselovsky et al. (2023) report that 33 to 46 percent of crowd workers utilize Large Language Models (LLMs) for labeling tasks, indicating a shift towards automation in data processing.

To ensure our categorization is reproducible we use the beta feature 'seed', which is a set seed parameter that aims to make the output of the model more deterministic.[1] This outputs a system fingerprint if the analysis is rerun holding the input text, the prompt, and the temperature constant as a change in system fingerprint indicates that there may have been a change in the model itself. To access this feature, we use the 'gpt-4-1106-preview' model which has the most consistent and reproducible output. This ensures that no observations are missed because of errors in the structured output, additionally, we set the script to try again if there is any error so that all comments are given a category in each iteration. We also set the temperature parameter equal to zero. This means that the most likely outputs are given a higher weighting in the softmax layer. Finally, we asked for the language models to provide reasoning and spreadsheets with this information so that it is possible to identify if the models use the same key information or whether they are focusing on different keywords in determining their category.

Below we provide the prompts used for categorization.

F.1 Prompt system for categorizing 'New lessons learned' answers

The Prompt

> *You are a highly detailed and analytical assistant.*
> *Your task is to categorize comments based on the following specific categories, which cover a wide range of topics from cultural festivals to economic challenges and educational barriers.*
> *The comment is in response to the question "If this is your village what would you suggest your neighbors to do?"*
> *Each category is outlined below along with a concise example that reflects the style of answers we're working with:*
>
> *Categories and Definitions:*
>
> - Education and Capacity Building (c1): Covers formal and informal education, training, and any activity aimed at improving knowledge, skills, and

[1] https://platform.openai.com/docs/api-reference/chat/create

competencies. Example: Conducting workshops on environmental education for community leaders.

- Livelihood and Artisanal Crafts (c2): Encompasses economic activities related to craftsmanship, artisan skills, and professions outside the agricultural sector, focusing on income generation and cultural heritage. Example: Organizing a local market for handmade pottery and ceramics.
- Cultural, Artistic, and Community Engagement (c3): Includes activities that promote cultural preservation, artistic expression, and community participation, strengthening social bonds and cultural identity. Example: Facilitating a community mural project that reflects local history and values.
- Agriculture and Rural Development (c4): Specifically targets agricultural activities, including farming and related rural development practices, emphasizing sustainability and environmental stewardship. Example: Initiating a community-supported agriculture (CSA) program to connect local farmers with urban consumers.

When analyzing an answer:

1. Identify any explicit mentions or clear implications related to the categories.
2. If the answer is general or lacks specific details, infer the most likely category based on broad definitions of keywords in the answer and categories.
3. Provide a reasoned analysis for your categorization, including how certain words or phrases led you to associate the answer with specific categories.
4. Look back at all the suggested categories if there are more than two, do the analysis again focusing on economic definitions of the keywords, and be more precise with categorization.
5. If there are still multiple categories, select only the most directly relevant.

For each answer, provide the following as a single YAML format:

- **comment**: (original comment)
- **category_index**: [list the category codes that apply, e.g., c1, c2, ...]
- **reasoning**: [provide a detailed explanation for each category chosen]

Your analysis should aim to capture the nuance and breadth of each comment's potential relevance to the categories, especially when direct information is limited.

F.2 Prompt system for categorizing 'What to do', 'Why', and 'How to encourage neighbors' answers

This prompt will take each answer and compare it to the categories. The best approach is to provide step by step instructions of what you want the tool to do, similar to how you would set up a data labelling task for a research assistant. This helps to ensure that the output is consistent for all answers. Where the prompt asks for YAML output this is a format that can be extracted directly into a dataframe. The bolded words such as **answer** and **keywords** refer to columns in the final dataframe that we would like to output. **category_index** returns an id for the category instead of rewriting the entire category.

The prompt

You are a highly detailed and analytical assistant.
Your task is to categorize comments based on the following specific categories, which cover a wide range of topics from cultural festivals to economic challenges and educational barriers.
Each category is outlined below along with a concise example that reflects the style of answers we're working with:

Enter the Categories here

When analyzing an answer:

1. Identify any explicit mentions or clear implications related to the categories.
2. If the answer is general or lacks specific details, infer the most likely category based on broad definitions of keywords in the answer and categories.
3. Provide a reasoned analysis for your categorization, including how certain words or phrases led you to associate the answer with specific categories.
4. Look back at all the suggested categories if there are more than two, do the analysis again focusing on economic definitions of the keywords, and be more precise with categorization.
5. If there are still multiple categories select the single most relevant category.

For each answer, provide the following as a single YAML format:

- **answer**: (original answer)
- **category_index**: [list the category codes that apply, e.g., c1, c2, ...]
- **keywords**: [list keywords used to inform the categorization]
- **definitions**: [list of keyword/key terms definitions e.g., keyword; definition]
- **reasoning**: [provide a detailed explanation for each category chosen]

Your analysis should aim to capture the nuance and breadth of each comment's potential relevance to the categories, especially when direct information is limited.

G Structured Ethics Appendix

1. Policy Equipoise

Ethical approval for this project was received from the Monash University Human Research Ethics Committee (ref no. 35152). The study compared a documentary about the lives of Santals to an active control documentary on flower farming of identical length. Both arms involved the same field procedures and compensation structure, ensuring policy equipoise and minimal incremental risk across conditions. A formal local IRB was not available for this project. Instead, we relied on Monash HREC oversight and collaboration with two Bangladeshi organizations (Ashrai and SARCH) to align procedures with local norms. All participants had an equal opportunity to be assigned to any study arm, and the allocation did not deprive any group of known superior benefits. Randomization was conducted at the village level, and within villages, household selection followed pre-specified random protocols.

2. Role of researchers with respect to implementation

Researchers were not "active" in the implementation. The intervention (screening the documentaries) and data collection were carried out by local NGOs (Ashrai for implementation; SARCH for survey work). The research team designed the study and regularly monitored compliance and made surprise visits to the field, but did not directly implement the intervention.

3. Potential harms to participants or nonparticipants from the interventions or policies

Informed consent was obtained for participation in all surveys and for the film screening, including for photography used exclusively for emotion-coding. Consent materials were delivered in Bangla; for participants with limited literacy, enumerators read the statement aloud, answered questions, and recorded verbal consent, with written consent obtained where feasible. Consent for photography was *separate* and refusal to be photographed did not affect eligibility, procedures, or compensation. The intervention (viewing a documentary film) was low-risk and voluntary, with no foreseeable physical, psychological, or economic harm for participating or not participating. Participants could decline or withdraw at any time. The content of the film was reviewed to avoid severe distress or incitement. No participants were denied access to standard benefits or future services. Protocols for privacy and confidentiality were in place.

Importantly, the intervention targeted only ethnic majority Bengalis. Data were also collected from ethnic minority Santals, but solely to assess whether the intervention generated positive or negative impacts on this group. Our findings indicate that there was a net benefit to ethnic minorities as a result of the intervention.

4. Potential harms to research participants or research staff from data collection (e.g., surveying, privacy, data management) or research protocols (e.g., random assignment)

Data collection followed established protocols to ensure privacy, confidentiality, and informed consent, including context-appropriate consent procedures. No risks beyond normal survey

participation were identified for participants or research staff. Participation involved standard surveys and behavioral experiments with minimal burden or discomfort. Data were de-identified and securely managed. Enumerators were trained to pause any interaction if a participant expressed discomfort, to offer a short break or termination without penalty, and to escalate concerns to a field supervisor.

5. Financial and reputational conflicts of interest. Do any of the researchers have financial conflicts of interest with regard to the results of the research? Do any of the researchers have potential reputational conflicts of interest?

None.

6. Intellectual freedom

No funder, implementer, or host institution held contractual rights to restrict analysis or publication. The research team retained full independence over analysis, interpretation, and write-up.

7. Feedback to participants or communities

We have not provided feedback to participants or communities yet. But we plan to disseminate key research findings to participating communities via our local collaborating NGO, Ashrai (e.g., short village meetings or a two-page brief), subject to financial and logistical feasibility. This complements academic dissemination.

CALL FOR NOMINATIONS

$300,000 Nemmers Prize in Economics

Northwestern University invites nominations for the Erwin Plein Nemmers Prize in Economics, to be awarded during the 2026–27 academic year. The prize pays the recipient $300,000. Recipients of the Nemmers Prize present lectures, participate in department seminars, and engage with Northwestern faculty and students in other scholarly activities.

Details about the prize and the nomination process can be found at **nemmers.northwestern.edu**. Candidacy for the Nemmers Prize is open to those with careers of outstanding achievement in their disciplines as demonstrated by major contributions to new knowledge or the development of significant new modes of analysis.Individuals of all nationalities and institutional affiliations are eligible except current or recent members of the Northwestern University faculty and past recipients of the Nemmers or Nobel Prize. **Nominations will be accepted until January 14, 2026.**

The Nemmers prizes are made possible by a generous gift to Northwestern University by the late Erwin Esser Nemmers and the late Frederic Esser Nemmers.

Northwestern

Nemmers Prizes • Office of the Provost • Northwestern University • Evanston, Illinois 60208
nemmers.northwestern.edu