Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India[†]

By Diva Dhar, Tarun Jain, and Seema Jayachandran*

This paper evaluates an intervention in India that engaged adolescent girls and boys in classroom discussions about gender equality for two years, aiming to reduce their support for societal norms that restrict women's and girls' opportunities. Using a randomized controlled trial, we find that the program made attitudes more supportive of gender equality by 0.18 standard deviations, or, equivalently, converted 16 percent of regressive attitudes. When we resurveyed study participants two years after the intervention had ended, the effects had persisted. The program also led to more gender-equal self-reported behavior, and we find weak evidence that it affected two revealed-preference measures. (*JEL* D63, D91, I21, J13, J16, 012)

Gender inequality exists in every society, but it is especially acute in many developing countries. Compared to men and boys, women and girls have fewer educational opportunities, less autonomy in decisions about their marriage and fertility, and more restrictions on their labor market participation and even their physical mobility and friendships (Duflo 2012; Jayachandran 2015). Many of these gender gaps have not narrowed in recent decades despite economic progress. Cultural norms often underpin these disparities, and economic development alone is unlikely to eliminate them (Alesina, Giuliano, and Nunn 2013; Jayachandran 2021).

Various policies might help foster greater gender equality in the face of restrictive gender norms. These include laws that guarantee equal rights, subsidies to encourage investment in girls, and programs that impart skills or target resources to women. A large literature has studied these approaches.

This paper focuses on a less common—and less commonly studied—approach: trying to directly change people's gender attitudes. We study an intervention that used discussion and persuasion to reduce participants' support for restrictive gender norms and to increase the value they place on equality. That is, it aimed to change their preferences.

Gender attitudes, even those rooted in centuries-old cultural norms, are amenable to change. For example, reserving seats for female politicians has helped curtail negative stereotypes about women as leaders in India (Beaman et al. 2009), and television shows have changed fertility preferences in multiple settings (Jensen and Oster 2009; La Ferrara, Chong, and Duryea 2012). The distinctiveness of the intervention we study is that reshaping gender attitudes was its primary goal.

The intervention, which we evaluate through a randomized controlled trial, was implemented in secondary schools in the state of Haryana, India. It engaged seventh- to tenth-graders in classroom discussions about gender equality, with a 45-minute session held every three weeks for two and one-half school years. The sessions taught facts and endorsed gender equality, and as importantly, prompted students to reflect on their own and society's views. Discussion topics included gender stereotypes, gender roles at home, girls' education, women's employment outside the home, and harassment. Some sessions taught communication skills to help students convey their views to others so that they can, for example, persuade their parents to permit them to marry at a later age. The program's messaging combined a human rights case for gender equity with pragmatic reasons to value women, such as their economic contributions.

Breakthrough, a nonprofit organization with extensive experience in gender-equality programming, designed and implemented the intervention. The Government of Haryana allowed Breakthrough to conduct the classes in schools as part of the regular school day. The government was interested in eroding some of the existing gender norms, and schools offer governments a powerful platform to shape the next generation's views. Secondary school students were specifically targeted because adolescence is a critical time in the development of morality and identity formation, with adolescents being young enough to still have malleable attitudes but mature enough to reflect on complex moral questions (Kohlberg 1976; Markus and Nurius 1986).[1]

Gender inequality is rife in India. While boys and girls start secondary school at the same rate, only 0.80 girls enroll in tertiary schooling for every boy (World Bank 2011). Early marriage is common, and many women have limited agency (Kishor and Gupta 2004). India's female labor force participation rate is among the lowest in the world (Field, Jayachandran, and Pande 2010; Klasen and Pieters 2015; Afridi, Dinkelman, and Mahajan 2018). It also has one of the most male-skewed sex ratios worldwide, due to the widespread practice of sex-selective abortion (Sen 1990). Among children age 0 to 6 years, there are 1.09 boys per girl; Haryana's sex ratio of 1.20 is the most male-skewed among Indian states (Jha et al. 2006; Government of India 2011).

---

[1] There were also practical reasons for targeting adolescents rather than younger students. The government was less comfortable with younger children discussing possibly sensitive topics like sexual harassment.

Our study encompasses 314 government secondary schools across four districts in Haryana. We collected data from roughly 14,000 students, both boys and girls. We analyze the intervention's effects a few months after the program ended and again two years later when the students were, on average, 17 years old.

The main outcome we study is participants' gender attitudes: that is, their views about what is right and wrong or desirable and undesirable, such as whether it is wrong for women to work outside the home and whether it would be good to have more women in politics. The second outcome is educational and career aspirations; for this outcome, we hypothesized that the intervention would only influence girls. The third outcome is self-reported behaviors influenced by gender norms. Only when the participants become adults can we assess impacts on major outcomes that they have control over, such as employment and childbearing. Nonetheless, examining day-to-day behavior in adolescence can provide insights on whether reshaping gender attitudes translates into behavior change. We focus on behaviors that adolescents likely have some say over, such as chores done at home and interaction with opposite-gender peers. One reason why attitude change might be insufficient for behavior change is people's desire to conform to social norms. A boy who believes that he should help out with chores might worry about the social sanctions he would face if he did so. For this reason, we examine perceptions of social norms as a secondary outcome.

We find that the intervention made gender attitudes more progressive by 0.18 standard deviations in the short run (three and one-half months after the program ended, or three years after baseline). The measure of attitudes is an index that aggregates several survey responses pertaining to support for gender equality. The effect size is equivalent to newfound support for gender equality in 16 percent of the cases where a student initially held a gender-regressive view.

What is especially striking is that these effects persisted. We continue to find a large effect on attitudes—0.16 standard deviations—in the medium run (two years after the program ended, or five years after baseline).

The program also influenced participants' self-reported behavior. An index of gender-equal behaviors increased by 0.20 to 0.23 standard deviations, in both the short run and medium run. In addition, we added two objective measures of behavior to the second end line. The first was a revealed preference measure of girls' educational intentions, namely whether they submitted a college scholarship application. The second measure tracked boys' and girls' likelihood of signing a public petition to end the dowry system. We find weak evidence of a treatment effect in the hypothesized direction for both behaviors. Finally, we find no evidence that the intervention increased girls' stated educational and career aspirations, which were quite high to begin with.

We investigate how the treatment effects vary based on two prespecified characteristics, student gender and parents' gender attitudes. We find no evidence of heterogeneity based on parents' attitudes, but we find important differences in treatment effects between boys and girls. At the first end line, the effect size on attitudes is somewhat larger for boys than girls, but we cannot statistically reject that the effect size is identical for the two groups. By the second end line, the effect on attitudes is significantly larger for boys. In addition, behavior change is significantly more pronounced among boys in both the short and medium run. For example, boys

report doing more chores, but girls do not report doing fewer, and only boys report an increase in how much they encourage their older sisters to pursue a college education. With chores, the smaller effect for girls can be explained by the asymmetry in the outcome. Greater gender equality maps to girls seeking a lower burden of chores, which others might not grant them, while boys can voluntarily help out more. But the fact that we also see heterogeneity by gender in support given to sisters suggests a broader phenomenon of girls facing more constraints on their behavior. This pattern highlights that, because behavior change requires not just the desire but also the ability to act differently, the very fact of boys' and men's greater power in society makes it important to include them in interventions aimed at increasing girls' and women's power.

A key concern with self-reported outcomes such as gender attitudes is the possibility of social desirability bias. The specific concern in our experiment is that there might be more social desirability bias in the treatment group, i.e., there could be experimenter demand effects. Participants in a program that explicitly tried to influence their support for gender equality might disingenuously express more gender-progressive views to present themselves in a good light to the surveyors. The persistence of the effects two years after the program ended provides some reassurance that it genuinely changed participants' views; it seems likely that experimenter demand effects would fade out with time. However, to address this concern more rigorously, we use the Marlowe-Crowne social desirability scale, which is a survey module developed by social psychologists to measure a person's propensity to give socially desirable answers (Crowne and Marlowe 1960). The module, which we included in the baseline survey, asks respondents if they have several too-good-to-be-true traits such as never being jealous of another person's good fortune and always being a good listener; those who report more of these traits are scored as having a higher propensity to give socially desirable answers.

We find that respondents with a high propensity for social desirability bias express more support for gender equality overall for the sample, but importantly, this pattern is not truer for the treatment group than the control group. In other words, the positive treatment effects on self-reported attitudes and behavior are similar in magnitude for respondents with a low versus high propensity for social desirability bias. We view this analysis as an important check on the validity of our results, and think that our approach could be useful in a wide array of studies in which experimenter demand is a concern. The method allows one to test for bias for any and all outcomes, so it complements techniques such as list experiments and revealed-preference measures, which often must focus on a narrower set of outcomes for logistical reasons.

Our study contributes to the literature on endogenous preferences, specifically on the formation of gender-related preferences.[2] Besides political quotas (Beaman et al. 2009) and television (Jensen and Oster 2009; La Ferrara, Chong, and Duryea 2012), other factors that have been shown to positively influence gender attitudes include mothers' employment (Fernandez, Fogli, and Olivetti 2004), having daughters or sisters (Washington 2008; Healy and Malhotra 2013), serving with women in the

---

[2]Recent work has also studied the formation of preferences toward different castes and toward children from poorer families in India (Lowe 2021; Rao 2019), Muslims in the United Kingdom (Alrababa'h et al. 2019), and racial minorities in the United States (Carrell, Hoekstra, and West 2019).

military (Dahl, Kotsadam, and Rooth 2021), having teachers who hold weaker gender stereotypes (Carlana 2019), and having female role models (Porter and Serra 2020).

Unlike most of the studies above on endogenous gender preferences, our research examines an intervention that intentionally changed preferences. It thus also sits within the literature on persuasion, or communication expressly designed to change preferences or beliefs (DellaVigna and Gentzkow 2010). Much of the economics literature on persuasion focuses on ways to influence consumer or political preferences. Closer to our work are studies across the social sciences on attitude change related to intimate partner violence (Gupta et al. 2013; Abramsky et al. 2014; Pulerwitz et al. 2015; Green, Wilke, and Cooper 2020), racial minorities (Donovan and Leivers 1993), immigrants (Hopkins, Sides, and Citrin 2019; Grigorieff, Roth, and Ubfal 2020), and women in STEM (Moss-Racusin et al. 2018), as well as studies that shift people's perceptions of social norms about gender or about ethnic discrimination and violence (Bursztyn, Gonzalez, and Yanagizawa-Drott 2020; Aloud et al. 2020; Paluck 2009). Our study is also related to Cantoni et al. (2017), which finds that Chinese students taught with textbooks designed to convey pro-Communist messages express more pro-government views and skepticism of free markets as adults.

We also add to the rapidly growing literature on educational/training interventions aimed at increasing girls' and women's agency and opportunities in developing countries. Related work includes Bandiera et al. (2020) on female empowerment and livelihood training in Uganda, Buchmann et al. (2018) on empowerment training and financial incentives to delay marriage in Bangladesh, Ashraf et al. (2020) on negotiation skills training for girls in Zambia, Edmonds, Feigenberg, and Leight (2020) on life skills training for girls in India, and McKelway (2020) on self-efficacy training in India. While most of these interventions primarily impart human capital to women and girls—either traditional skills or positive psychological traits—the central (though not exclusive) aim of the intervention we evaluate is to influence participants' preferences, specifically their attitudes about traditional gender roles. Our study is also relatively unusual in the literature on women's empowerment in that it focuses on both boys and girls.

I. Description of the Intervention

This project emerged from the Government of Haryana's interest in testing policies to narrow gender gaps in its society. Breakthrough, a human rights organization specializing in social change campaigns related to gender, designed an intervention aimed at changing adolescent boys' and girls' views about gender norms, and implemented it in government schools with the state government's permission.

The participants in the program were the cohorts in grades 7 and 8 in the academic year 2014–2015 when the program launched. It ran from April 2014 to October 2016, so one cohort participated in the program in grades 7, 8, and half of grade 9, and the other in grades 8, 9, and half of 10. Grades 7 to 10 have high enrollment and low dropout in Haryana, so the program could reach a large share of the underlying age cohorts and have limited attrition due to school dropout (significant dropout occurs after grade 10) (DISE 2011).

The objective of the program, which was named *Taaron ki Toli*, or Legion of Stars, was to create awareness of gender-based discrimination, change dominant gendered perceptions, promote gender-equitable attitudes, raise girls' aspirations, and provide tools to participants to translate attitude change and greater aspirations into behavior change. The program aimed to ultimately influence a wide range of behaviors related to female education, mobility, work, marriage, and fertility, for both female participants and male participants' female family members (e.g., their future wives).

The program emphasized both human rights and pragmatic (i.e., instrumental) reasons for giving girls and women more opportunities. For example, it conveyed that equal opportunity for education is a universal human right. The hypothesis is that this rights-based message would increase how much participants value girls having access to higher education; a boy would get disutility from seeing his sister denied the same chance to attend college that he has. As an example of a pragmatic argument for girls' education, the intervention informed participants that outcomes for children improve when their mother is more educated. The hypothesis is that this information causes updating of beliefs; girls want to stay in school longer, and both boys and girls will want to educate their daughters down the road.

To ensure that the intervention would be widely accepted, Breakthrough engaged with multiple stakeholders at the state, district, and sub-district levels, gathering input from education officials, school principals, and teachers as they developed the program. This helped them design a program that might later be integrated into the standard school curriculum.[3]

The program featured classroom sessions, each 45 minutes long, led by a Breakthrough facilitator. The regular teacher was welcome to stay for the sessions or leave the classroom and have a break. There were a total of 27 sessions spread over two and half years.[4] Breakthrough hired 15 facilitators, 13 of whom were male, to cover the 150 treatment schools.[5] The facilitator visited each school roughly once every three weeks. Other elements of the program included a one-time training for one teacher per school, optional youth clubs, and school-wide activities such as street theater performances held about once a year.[6]

---

[3] Scale-up options include hiring special-purpose teachers to lead the sessions, each of whom covers multiple schools; having regular school teachers deliver the lessons; incorporating some of the content into textbooks; or continuing to have NGOs implement the program. The Government of Punjab (India) announced in March 2021 that it planned to roll out the program in all of its upper primary schools in late 2021, to be taught by regular social studies and moral science teachers (Menon 2021).

[4] The total dosage was 20 hours, or about 23 hours adding in the school-wide assemblies. As comparisons, the negotiation program for girls in Zambia studied by Ashraf et al. (2020) consisted of six two-hour sessions (12 hours total); the safe space groups in Bangladesh evaluated by Buchmann et al. (2018) met for about 200 hours total over six months; and the empowerment and livelihood clubs in Uganda evaluated by Bandiera et al. (2020) were open five afternoons per week for two years (over 500 hours).

[5] The majority male staff was due to the very restrictions on women's mobility that the program aimed to change; the job entailed traveling across a geographic area spanning 10 treatment schools. We lack the statistical power to shed light on heterogeneous effects by facilitator gender. For both boys and girls, messages about gender equality could be more powerful coming from a man, or conversely from a woman. It is also plausible that students are most responsive to messages from same-gender instructors.

[6] This curriculum was delivered to the two study cohorts, finishing in the middle of the 2016–2017 school year. Breakthrough received additional funding and resumed activities in 59 of the 150 treatment schools in 2017–2018, offering an extra module to the younger of our two study cohorts. This added 0.2 years of dosage on average (59/150 * 52 of sample in younger cohort * 1 year), or 2.1 additional sessions per participant. Breakthrough also initiated the full curriculum with new cohorts, 2 to 5 years younger than our study cohorts, in these schools. We do not expect spillovers from these younger children to our study participants to have added much dosage, especially

The classroom sessions were discussion-based, with more student participation than is typical in secondary schools classes. Perhaps for this reason, and also because having a dynamic teaching style was one of the hiring criteria for facilitators, student reaction to the program was generally very positive. To complement the in-class material and encourage further reflection, the facilitators assigned some homework assignments such as to write stories and record observations, and they encouraged students to talk to their family members about what they were learning. The sessions did not displace one specific subject like math or history; they crowded out a roughly even mix of material in other subjects. While the program could have harmed other learning by displacing instructional time, the discuss-and-debate style could have strengthened students' critical thinking and speaking skills, conferring benefits beyond the gender focus.

Discussion topics for the sessions included gender identity, values, aspirations, gender roles and stereotypes, and recognition and tolerance of discrimination. For example, one session focused on household chores. Students broke out into groups and listed whether males or females did various chores in their households. They then reconvened and discussed the answers. When the pattern emerged that women and girls did most of the chores, the facilitator asked why that was and whether it was fair. The class discussed why women cook at home, but men are cooks in restaurants, with the latter role earning more status in society. A few of the sessions aimed to impart skills such as public speaking, communication between the genders, and leadership, which could enable gender-equitable attitudes to translate into behavioral change. For instance, girls might be able to negotiate greater independence with their parents, leading to more freedom of movement in the short run and greater occupational choice in the long run. Through this curriculum, students explored gender identity and stereotypes, gained a better understanding of gender inequities and their consequences, understood their rights, and were encouraged to communicate and act on what they had learned. To map this to standard concepts used in economics, the intervention aimed to change students' preferences (i.e., their moral views on gender inequality); their factual beliefs (e.g., greater realization that restricting women's employment leaves money on the table); and their skills (e.g., how to persuade their parents to let them go to college).

II. Study Design and Data

A. *Experimental Design*

We conducted a randomized evaluation of the gender attitude change program in a sample of 314 government schools across Sonipat, Panipat, Rohtak, and Jhajjar districts in the state of Haryana, India.[7] The unit of randomization was the school.

---

since most of the older cohort had switched to a new school for grade 11 by then. We became aware of these additional activities in 2019.

[7] The government prioritized these districts based on their skewed sex ratios. The 2011 child sex ratio was 1.25 in Sonipat, 1.19 in Panipat, 1.22 in Rohtak, and 1.28 in Jhajjar. These districts do not necessarily have more male-biased preferences than the rest of Haryana or north India, but their low fertility rate (they are near New Delhi) means that son preference translates into a higher rate of sex-selective abortions to ensure having at least one son in the family (Jayachandran 2017).
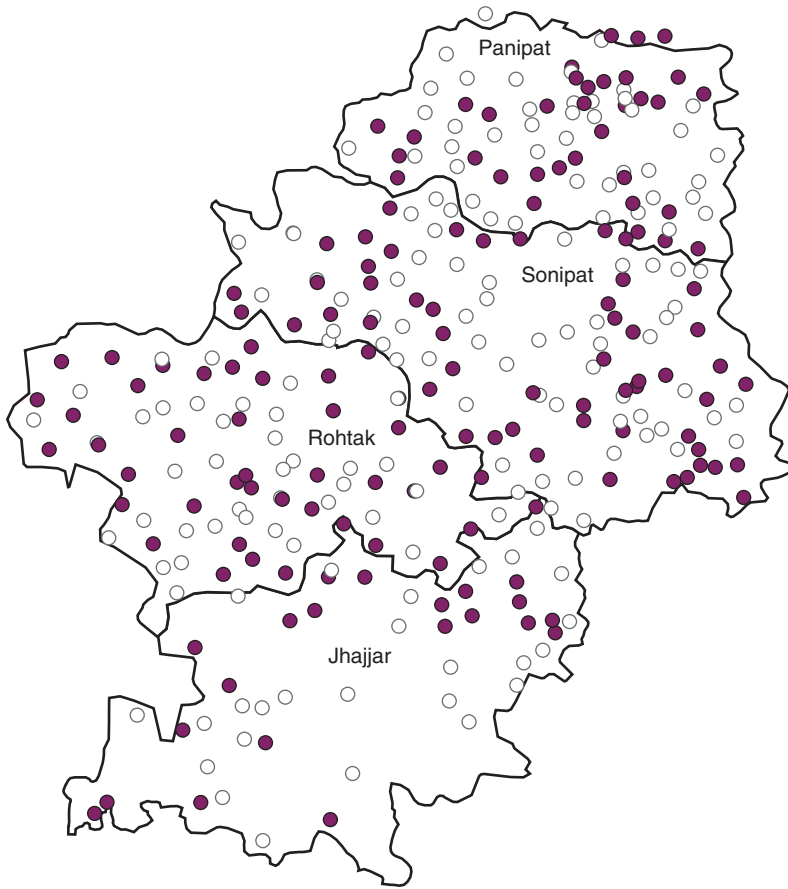
FIGURE 1. MAP OF TREATMENT AND CONTROL SCHOOLS WITHIN THE STUDY DISTRICTS

The sample size of 314 schools was chosen to be able to measure the short- and medium-run effects of the program on gender attitudes, aspirations, and behavior, as well as long-term effects on educational attainment, occupational choice, marriage, and fertility that might emerge up to ten years after the program ended. There were 607 government-run secondary schools that offered grades 6 through 10 across the four districts. We first restricted attention to the 346 schools that officially enrolled at least 40 students in grades 6 and 7 combined, and then eliminated schools with low actual enrollment based on a preliminary visit. In cases where a village had more than one government secondary school, we chose at most one of them for the sample to minimize the possibility of spillovers. Of the 314 schools in the sample, 59 enroll only girls, 40 enroll only boys, and the remaining 215 are co-ed. Official enrollment in these schools averages about 80 students per grade (DISE 2011).

We randomly selected 150 of the sample schools to be in the treatment group; the remaining 164 serve as control schools. Figure 1 shows the four study districts and the schools assigned to the treatment and control groups. The randomization was stratified by district, co-ed status of the school, school size, and distance to the district headquarters. Table 1 reports baseline characteristics of schools by treatment

Table 1—Descriptive Statistics: School and Student Characteristics at Baseline

Variable	Treatment	Control	Standardized diff
Number of schools	149	164	
Urban	0.107	0.073	0.119
	[0.311]	[0.261]	
School is co-ed	0.698	0.677	0.045
	[0.461]	[0.469]	
Males in grades 6 and 7	66.427	65.270	0.028
	[45.948]	[35.963]	
Females in grades 6 and 7	75.125	74.212	0.015
	[60.081]	[58.344]	
Number of students	7,051	7,758	
Student's age	11.833	11.854	−0.017
	[1.261]	[1.250]	
Female	0.566	0.544	0.044
	[0.496]	[0.498]	
Hindu	0.945	0.953	−0.036
	[0.227]	[0.211]	
Enrolled in grade 6	0.526	0.521	0.011
	[0.499]	[0.500]	
Scheduled caste	0.268	0.285	−0.039
	[0.443]	[0.451]	
Mother's age	35.462	35.572	−0.017
	[6.351]	[6.513]	
Father's age	40.497	40.611	−0.016
	[6.895]	[7.131]	
Mother is illiterate	0.370	0.374	−0.009
	[0.483]	[0.484]	
Mother works full-time	0.292	0.292	−0.002
	[0.455]	[0.455]	
Dwelling has flush toilet	0.155	0.130	0.070
	[0.362]	[0.337]	
Gender attitudes index	0.032	0.000	0.032
	[1.008]	[1.000]	
Girls' aspirations index	0.039	0.000	0.039
	[1.001]	[1.000]	
Self-reported behavior index	−0.012	0.000	−0.012
	[0.994]	[1.000]	
Social desirability score	−0.025	0.000	−0.025
	[1.018]	[1.000]	
High social desirability score	0.374	0.375	−0.002
	[0.484]	[0.484]	

Note: F-stat for joint significance of above baseline student variables is 0.924.

status. The first panel confirms that the two samples are balanced on various school characteristics such as urban/rural and number of male and female students.

B. *Enrollment of Study Participants and Baseline Data Collection*

The baseline survey was conducted between August 2013 and January 2014, covering 14,809 students. The data collection was conducted by the Abdul Latif Jameel Poverty Action Lab, South Asia. All students (as well as parents, principals, field staff, etc.) were blind to treatment status when recruitment for the study and baseline data collection took place.

To select students for the sample, we visited the schools and distributed parental consent forms to all sixth and seventh-graders who were present. These students would be in grades 7 and 8 in April 2014 when the program began at the start of the next school year. The parental consent rate was high; 84 percent of forms were returned. Perhaps surprisingly, the rate does not vary by gender or by village-level proxies for gender norms, such as the child sex ratio and female employment rate from the 2011 census. Anecdotally, lack of consent was usually due to the student losing or forgetting the form.[8]

We randomly chose the study participants from among those whose parent gave consent, with a target of 45 students per school, stratified by gender and grade with a ratio of 3:2:2:2 for female sixth:male sixth:female seventh:male seventh. We included more girls than boys in the sample because there are more girls enrolled in government secondary schools, and we sampled more grade 6 girls than grade 7 girls because we expected lower attrition for younger grades.[9] Students also needed to personally assent to participating in the study and be present at school on the baseline survey day. The 35-minute baseline survey took place on the school premises.

We mistakenly omitted one school from the baseline survey. This school was randomized into the treatment group, and it received the intervention. We collected end line data in the school and include it in the analysis, imputing baseline variables with the gender-specific sample average for the district.[10]

In addition, we surveyed one parent for a random 40 percent subsample of the students ($N = 6,022$). We selected at random whether to interview the father or mother. We interviewed the parents at home, which added survey costs and is why we did not include all parents. We construct a gender attitude index for the parent based on nine attitude questions, which we use to understand how parental attitudes influence program impacts.

Table 1 summarizes baseline characteristics of the sample. The boys and girls were about 12 years old, on average. Religious and caste variables line up with the overall demographics for these districts, as reported in the census (Government of India 2011); the participants are predominantly Hindu. Mothers' average age was 36 years and fathers' was 41 years. There is a high illiteracy rate for mothers, reflecting the low level of female schooling in the parents' generation. Consistent with India's low female labor force participation rate, only 29 percent of mothers are employed full-time. Baseline variables are balanced between the treatment and control groups. An F-test of joint significance fails to reject balance between the study arms. Online Appendix Table 1 shows summary statistics separately for girls and boys, which are also balanced between the treatment and control groups.

In the baseline survey, we included the Marlowe-Crowne module, designed by social psychologists to measure a person's propensity to give socially desirable responses (Crowne and Marlowe 1960). The module asks the respondent whether

---

[8] Parents had to consent to their child participating in the study, but the program was added to the curriculum in treatment schools by the government, which did not offer parents a way to opt out. The classes were not held on a fixed day, so keeping a child home to avoid the classes would not have been straightforward. Anecdotally, parental complaints about the program to schools and the education department were negligible.

[9] Parents are more likely to send their sons than daughters to private schools. Because wealthier families use private schools, the boys in government schools are, thus, from poorer families than the girls, on average.

[10] We distributed consent forms to students present during a school visit just before the end line survey and then randomly chose sample students from among those with parental consent who assented to participate.

he or she has certain almost saintly personality traits (e.g., "I am never irritated by people who ask favors of me"). Because the module is designed to use traits that people are unlikely to truly have, the interpretation when someone reports having more of these traits is that she has a stronger concern for social approval. A caveat is that some of the variation might reflect actual differences in possessing these desirable traits. The module was developed in the United States but has been validated in several developing countries, including India (Mukherjee 1967; Vu et al. 2011). We used a 13-item version of the original 33-item module (Reynolds 1982). The questions are listed in the online Appendix. We combine the responses into an index, or social desirability score, which we use to investigate whether the treatment effect estimates are biased upward by experimenter demand effects. The social desirability score is balanced between the treatment and control groups, and its distribution by gender is shown in online Appendix Figure 1.

C. *End Line Data Collection*

We conducted a first end line survey shortly after the program ended (three and one-half months, on average), or about three years after the baseline survey. We then conducted a second end line two years later.

Data collection for the first end line survey occurred between November 2016 and April 2017. We resurveyed 13,943 of the 14,809 students surveyed at baseline, which corresponds to an attrition rate of 5.8 percent. The end line sample also includes an additional 44 students from the sample school that we mistakenly did not survey at baseline, yielding a total sample for the first end line of 13,987 students.

Online Appendix Table 2 shows that sample attrition does not differ significantly between the treatment and control groups for either boys or girls. In addition, attrition in the treatment versus control group is not differential by baseline outcomes.[11] Three fourths of students were surveyed at school at the first end line. Several students had moved to a different school, either in the same or a different village, or dropped out of school. These students were surveyed at home. If the student had moved to another village that was far from the survey districts, we conducted a truncated phone survey (0.1 percent of respondents).[12]

We conducted a second end line survey between January and July 2019, which was two to two and one-half years after the intervention had ended and when the students were finishing or had just finished grades 11 and 12 (if they had not repeated a grade). The survey was conducted in students' homes, rather than at schools; the highest grade offered in most of the sample schools was grade 10, so the participants were scattered across various schools (or had dropped out). We again conducted truncated phone surveys for the small share of respondents who had moved. The attrition rate was 7.9 percent, and attrition is not significantly correlated with

---

[11] Online Appendix Table 3 details the reasons for attrition, which include permanent or long-term migration, death or poor health, refusal to participate by the student or parent, not being available at the time of their appointment, and inability to track the respondent.

[12] Online Appendix Table 2 shows that treatment status is not significantly correlated with the survey location. Online Appendix Table 3 summarizes participants' schooling status at end line: 86 percent of girls and 76 percent of boys were enrolled in the same school as baseline; boys are more likely to have switched to a private school and also to have dropped out of school. The table also reports that over 85 percent of the treatment group was aware of the program activities.

treatment status or correlated with baseline characteristics differentially by treatment status. The sample size for the second end line is 13,685 individuals.

D. *Primary Outcomes: Attitudes, Aspirations, and Behavior*

We prespecified three primary outcomes for the first end line: gender attitudes, girls' aspirations, and self-reported gender-related behavior.[13]

We combine 17 gender attitude variables into a variance-weighted index, following Anderson (2008). The specific variables and procedure for constructing the index were prespecified (see the online data Appendix for more details). We measure gender attitudes mostly through direct questions about female and male roles and rights (e.g., whether women should work outside the home, the appropriate age of marriage for girls), plus a vignette about investing in a son's or daughter's education.

Gender attitudes are balanced between the treatment and control group at baseline (see Table 1).[14] They are also quite regressive. For example, about 80 percent of boys and 60 percent of girls believe that a woman's most important role is being a good homemaker. This pattern that girls are less likely than boys to endorse gender-discriminatory positions is seen for each of the attitude questions (see online Appendix Table 4).

We measure girls' aspirations with a variance-weighted index combining 5 questions about educational and career aspirations. Girls have high aspirations relative to the actual rate of female college completion and employment in their communities. For example, at baseline, 72 percent of girls (compared to 77 percent of boys) expect to be employed and holding a white-collar job at age 25. This rate is much higher than the current employment rate of young women in India.

Arguably the most challenging outcome to measure was self-reported behavior. We focused on behaviors that are influenced by gender attitudes and gender norms, and that we expected adolescents to have some say over. We construct an index of 6 questions, asked of both boys and girls. Most of these questions are coded the same way for boys and girls (e.g., being supportive of sisters' career aspirations, level of interaction with the opposite gender), whereas household chores is coded so that, for girls, more gender-equal behavior is to do fewer chores whereas, for boys, it is to do more. We also included some questions applicable to only girls (e.g., mobility), which we do not include in the main index but examine in auxiliary analyses.

For the second end line, we again prespecified gender attitudes, girls' aspirations, and behavior as primary outcomes. We construct the attitudes index identically between the first and second end lines, using the same questions and weights (which are based on the first end line data). This makes the magnitude of the effect directly comparable across the two waves. For aspirations and behavior, we updated

---

[13] The preanalysis plan (PaP) for the first end line survey was posted to the AEA RCT Registry at the beginning of end line data collection in November 2016. It specified the primary outcomes and how they would be constructed, the secondary outcomes, heterogeneity analyses, and the procedure for choosing control variables. The PaP for the second end line was posted in December 2018 and specified similar information. The two PaPs and a short document that lists the ways in which we deviated from the PaPs are available at bit.ly/PaP4RAGA.

[14] The gender attitudes module at baseline was shorter than the end line module. The baseline attitude index aggregates nine variables.

the modules, as the relevant questions changed as the sample grew older, so we construct the variance-weighted indices independently for the two end lines.

We added two revealed-preference measures as primary outcomes in the second end line. The first captures girls' intent to pursue a college education. We set up a girls' scholarship program for college expenses and use as an outcome whether respondents filled out and mailed in the application, which we gave to them at the end of the survey visit and which required some time to fill out and submit. Application submission serves as an "intensive margin" measure of how intensely a girl aspires to attend college combined with her expectation that she will be allowed to attend college. The second revealed-preference measure focuses on both boys' and girls' willingness to publicly espouse a feminist position. We informed respondents about a petition to end the dowry system, with the names of signatories to be published in the local newspaper (through an advertisement we placed). The dowry system was not explicitly covered in the curriculum, though the topic likely came up in some discussions. Students were given a toll-free phone number to call to add their name to the petition. We use signing the petition as an outcome.

III. Empirical Specification

We hypothesized that the intervention would make participants' attitudes less discriminatory against females, raise girls' aspirations, and increase gender-equitable behavior. This section describes the estimation strategy used to test these hypotheses.

We estimate the following ordinary least squares regression, with one observation per student:

$$(1) \qquad Y_{ij} = \beta_0 + \beta_1 \, Treated_j + \beta_2 Y_{ij}^0 + \beta_3 \mathbf{X}_{ij} + \epsilon_{ij}.$$

Y_{ij} is the outcome variable measured at end line for student i in school j. $Treated_j$ is a binary variable that equals 1 if the school was assigned to the treatment group, and 0 otherwise. Thus, β_1 represents the average effect of the intervention on the outcome. The outcomes are constructed so that a higher value represents more gender progressiveness, so the hypothesis is $\beta_1 > 0$.

We control for Y_{ij}^0, the baseline analog of the outcome. The vector $\mathbf{X}_{ij}$ comprises other control variables, specifically grade-gender and district-gender fixed effects. When the outcome is an index, we include a missing flag for each component of the index.[15] We allow the error term, ϵ_{ij}, to be clustered at the school level, which is the level of randomization.

In addition, we estimate an alternative specification in which the regressors besides $Treated_j$ are chosen using the double LASSO procedure of Belloni, Chernozhukov, and Hansen (2014). Online Appendix Table 5 lists the control variables selected for each outcome and the larger set of potential controls from which the double LASSO procedure chose them.

We also test for heterogeneous treatment effects along prespecified dimensions: student gender and parents' attitudes at the first end line, and only student gender

---

[15] When a component of the index is missing for an observation, we construct the index using the remaining nonmissing variables. When the baseline outcome Y_{ij}^0 is missing, we impute its value with the district-gender mean.

at the second end line. Given the parsimonious set of primary outcomes and heterogeneity analyses, we do not adjust the statistical inference for multiple hypothesis testing.

IV. Short-Run Results

This section presents the results from the first end line survey, which was conducted three and one-half months after the intervention ended.

A. *Effect on Gender Attitudes*

Our first main finding is that the intervention made gender attitudes more progressive. Students in treatment schools have a 0.18 standard deviation higher attitude index than those in control schools ($p < 0.01$), as reported in Table 2, column 1.

One benchmark for the effect size is that end line gender attitudes are 0.50 standard deviations higher for girls than boys in the control group, as reported in online Appendix Table 6. Thus, the treatment effect is 36 percent as large as the status quo gender gap in attitudes. Also, a 1 standard deviation increase in parent gender attitudes is associated with student gender attitudes being 0.05 standard deviations higher; the treatment effect is much larger than this.[16]

Another way to express the effect size is that the intervention's "persuasion rate" was 15.5 percent (DellaVigna and Gentzkow 2010). In other words, the intervention succeeded in converting 15.5 percent of gender-regressive views into support for gender equality. The persuasion rate is calculated by stacking all of the variables in the attitudes index, which are coded as binary values. In the treatment group, on average 29.9 percent of views are gender-regressive, compared to 35.4 percent in the control group.

When we decompose the attitude index into thematic sub-indices, we find that the program had the strongest effect on attitudes about employment and other equal rights for women and girls, followed by education attitudes, as reported in online Appendix Table 7. The effect on gender-equitable fertility attitudes among girls is statistically significant but small (2 percentage points more likely to have gender-equal fertility preferences) and negligible and insignificant for boys; the intervention included very little discussion about fertility, in part because the government requested no discussion of sexual activity. The effects for each of the 17 variables that comprise the overall gender attitudes index are reported in online Appendix Table 8.

The results described above use our main specification, with the basic set of controls. The effect of the intervention on the attitudes index is very similar (0.17 standard deviations; $p < 0.01$) when we instead select control variables with double LASSO, as shown in online Appendix Table 9. To account for the possibility that attrition is endogenous to treatment, we also estimate Lee bounds on the treatment

---

[16] In Dhar, Jain, and Jayachandran (2019), we present an arguably better version of this parental attitudes benchmark, using indices constructed from the same set of questions for parents and students and collected at baseline for both groups. A 1 standard deviation increase in a parent's attitudes is then associated with a 0.11 standard deviation increase in the child's attitudes.

TABLE 2—TREATMENT EFFECTS ON ATTITUDES, ASPIRATIONS, AND BEHAVIOR (END LINE 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.180	0.030	0.196
	[0.020]	[0.024]	[0.021]
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	13,987	7,767	13,974

Notes: All regressions control for the baseline analog of the outcome, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

effects (Lee 2009). The attrition-adjusted lower bound on the point estimate is 0.16, as shown in online Appendix Table 10.

B. *Effects on Girls' Aspirations and on Self-Reported Behavior*

Turning to our second primary outcome, Table 2, column 2 shows that the program did not affect girls' aspirations. The average effect is 0.03 standard deviations and not significantly different from zero. Girls' aspirations were high to begin with. In addition, our measure does not capture how intensely the respondent held her aspirations, for example, how hard she would fight to be able to have a career. For these reasons, there might have been limited room for the intervention to raise the aspirations measure further.

The third primary outcome is self-reported behaviors influenced by gender attitudes. Behavior became more aligned with gender-progressive norms by 0.20 standard deviations ($p < 0.01$), as reported in Table 2, column 3. The magnitude and significance of this effect are robust to including additional control variables selected using double LASSO. Decomposing the behavior measure into sub-indices, we find that the intervention generated more interaction with the opposite sex for both boys and girls (see online Appendix Table 11). It also increased boys' participation in household chores—a shift in the direction of a more gender-equal division—and their support for their female relatives' ambitions. In addition, among girls, the program led to greater mobility (e.g., walking to school alone) but had no impact on decision-making power.[17]

C. *Assessing Bias Due to Experimenter Demand Effects*

An important concern when interpreting changes in self-reported outcomes is that participating in the program might have made salient what the socially desirable responses to our survey questions were without changing actual views. The treatment group likely became more aware that many outsiders to their community regard support for gender equality as laudable. If the program caused participants to refrain from making gender-discriminatory statements in their daily lives even

---

[17] Online Appendix Tables 12 and 13 report the effects for the component variables that comprise the aspirations and behavior indices.

Table 3—Robustness Check for Social Desirability Bias (End Line 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.190 [0.024]	0.018 [0.029]	0.196 [0.023]
High social desirability (Soc. D) score	0.106 [0.020]	0.062 [0.030]	0.060 [0.019]
Treated × High Soc. D score	−0.024 [0.030]	0.032 [0.043]	0.001 [0.028]
p-value: Treated + Treated × High Soc. D= 0	0.000	0.171	0.000
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	13,987	7,767	13,974

Notes: The social desirability (Soc. D) score is a baseline measure of the student's propensity to give socially desirable answers. High Soc. D score refers to having an above-median score among students. All columns control for the baseline analog of the outcome variable, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

without changing their deeply held beliefs, that would still be a true program benefit. The concern is if the estimates reflect what participants are willing to say to the surveyors, in ways not reflective of how they act outside the study environment.

To investigate whether this type of experimenter demand effect is upward biasing the estimated program impacts, we construct a social desirability score for each respondent using the Marlowe-Crowne module administered at baseline. The score measures a person's general tendency to present herself in a socially desirable way in the survey. We test for heterogeneous treatment effects based on the social desirability score. The worrisome pattern would be if the treatment effects were driven by students with a high propensity to disingenuously give socially desirable answers and vanished for those with a low such tendency.

As shown in Table 3, the main effect of having an above-median social desirability score is positive and significant for all three primary outcomes, suggesting some upward shading of responses overall for the sample. We view this pattern as reassuring, a validation that the score is successful in capturing a respondent's propensity to shade responses. Importantly, there is no more of this shading up in the treatment group than in the control group; the interaction terms are small and insignificant. For example, the 0.19 main effect of *Treated* in column 1 implies that the program's effect on attitudes is large and significant when we focus on the subsample with lower susceptibility to experimenter demand effects. Online Appendix Table 14 shows that we find similar patterns if we use the continuous measure of the social desirability score instead of an indicator for an above-median score.

Assuming that the Marlowe-Crowne score is primarily measuring a person's propensity to give socially desirable answers (as opposed to his or her saintliness), these results help allay the concern that the estimates reflect experimenter demand effects.

TABLE 4—GENDER-SPECIFIC TREATMENT EFFECTS ON ATTITUDES, ASPIRATIONS, AND
BEHAVIOR (END LINE 1)

	Gender attitudes index		Self-reported behavior index	
	Girls (1)	Boys (2)	Girls (3)	Boys (4)
Treated	0.161 [0.025]	0.204 [0.029]	0.142 [0.026]	0.260 [0.029]
Control group mean	0.237	−0.283	−0.086	0.102
Basic controls	Yes	Yes	Yes	Yes
Number of students	7,802	6,185	7,794	6,180
p-value: Girls = Boys	0.267		0.001	

Notes: All regressions control for the baseline analog of the outcome, grade and district fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

D. *Heterogeneous Treatment Effects by Student Gender and Parental Attitudes*

For several reasons, the program could have had different effects on boys and girls. Girls might have felt more invested in the program's messages. Alternatively, the ideas presented could have been more eye-opening for boys, and they might have related more to the mostly male facilitators. We thus analyze the effects of the intervention separately for girls and boys. The results are reported in Table 4.

For gender attitudes, while the point estimate for the treatment effect is somewhat smaller for girls than boys, we cannot reject that the program had the same impact for both genders. The statistical comparison of the two effect sizes is based on estimating the pooled regression that maps to the two panels of Table 4 (i.e., every regressor is interacted with *female*). The fact that girls started out more progressive than boys raises the issue of whether estimates of heterogeneity by gender are entangled with heterogeneity by initial attitudes. As shown in online Appendix Table 15, the patterns of gender heterogeneity are similar when we simultaneously allow for heterogeneity by baseline attitudes.[18]

For behavior, we find that the program had a significant positive impact for each gender, but a smaller impact for girls than boys. The p-value of this difference is < 0.01. One interpretation of this finding is that boys and girls can adopt gender-equal attitudes with relatively equal ease, but girls face more familial and societal constraints on translating their attitudes into behavior.

In light of the notable heterogeneity by gender, in subsequent tables we show the results separately by gender (and report inference based on interacted models). In addition, online Appendix Tables 9, 17, and 18 show our main robustness checks (i.e., double-LASSO-selected controls, Lee bounds, social desirability bias) separately for girls and boys.

The second dimension of heterogeneity that we prespecified was parental attitudes. In principle, the program could have had either larger or smaller effects for

[18] The gender heterogeneity analysis is also robust to correcting for the gap in economic status between girls and boys in government schools, which is due to boys' higher rate of attending private schools as discussed in footnote 9. See online Appendix Table 16, which controls for several wealth proxies in parallel to gender.

Table 5—Heterogeneous Effects by Parent Attitudes (End Line 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.174	0.054	0.179
	[0.027]	[0.033]	[0.026]
Treated × baseline parent attitudes	0.026	0.000	−0.039
	[0.025]	[0.026]	[0.021]
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	5,718	3,231	5,717

Notes: All regressions control for the baseline analog of the outcome, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

students whose home environment was more conservative. Table 5 reports this analysis, which uses the subsample for which we surveyed one of the parents at baseline. The index of parental attitudes is normalized to have a standard deviation of 1. When gender attitudes and girls' aspirations are the outcomes, the point estimates for the interaction coefficients are small compared to the main effects and insignificant (columns 1 and 2). Column 3 shows some weak evidence that the intervention engendered less behavior change among students from more progressive families, but the magnitude of this heterogeneity is small; it corresponds to the effect size differing by 0.04 standard deviations between the subsamples with above- and below-median parental attitudes (see online Appendix Table 19 for the results using a binary measure of parental attitudes). Overall, there is limited evidence that parental support for gender equality either facilitates or hinders the average success of the intervention.[19]

E. *Effects on Secondary Outcomes*

In addition to our primary outcomes, we prespecified a handful of secondary outcomes. One of them is perceptions of social norms. While the program only directly reached 100 to 200 adolescents per village, program participants might have started regarding pro-equality views as more mainstream and, thus, updated their belief about how common those views were in their community. Past research suggests that signals from institutions (Breakthrough in this case) can be effective in changing subjective perceptions of norms (Tankard and Paluck 2016).[20]

We examine parallel questions about (i) personally holding a positive gender attitude, (ii) believing one's community has a positive gender norm in that domain, and (iii) personally holding the positive attitude and believing the community will not oppose you if you act on it. We developed these questions by drawing on Bicchieri (2016). In addition to using these questions to assess how perceived social norms

---

[19] In exploratory analyses, we find no heterogeneous effects by whether the school was co-ed, facilitator gender, siblings' sex composition, the sex ratio in the village, or the female employment rate in the village.

[20] Recent evidence from India and Saudi Arabia suggests that people overestimate their community's opposition to female employment (Bernhardt et al. 2018; Burztyn, Gonzalez, and Yanagizawa-Drott 2020).

TABLE 6—TREATMENT EFFECTS ON PERCEPTIONS OF SOCIAL NORMS (END LINE 1)

	Social norms toward work			Social norms toward education		
	Student agrees:			*Student agrees:*		
	Women should be allowed to work (1)	Community thinks women should be allowed to work (2)	Women should be allowed to work and thinks community will not oppose them (3)	Women should be allowed to study in college even if it is far away (4)	Community thinks women should be allowed to study in college even if it is far away (5)	Women should be allowed to study in college and thinks community will not oppose them (6)
Panel A. Girls						
Treated	0.083	0.028	0.040	0.038	0.015	0.015
	[0.011]	[0.017]	[0.016]	[0.008]	[0.018]	[0.017]
Control group mean	0.848	0.518	0.587	0.935	0.623	0.695
Number of students	3,874	3,661	3,625	3,900	3,737	3,717
Panel B. Boys						
Treated	0.196	0.085	0.120	0.145	0.102	0.129
	[0.020]	[0.020]	[0.020]	[0.016]	[0.019]	[0.020]
Control group mean	0.496	0.337	0.316	0.758	0.557	0.571
Number of students	2,863	2,691	2,672	2,995	2,847	2,833
Panel C. Girls = Boys p-value						
	0.000	0.025	0.003	0.000	0.001	0.000

Notes: All columns control for grade and district fixed effects. Each respondent was given either the set of questions on norms about work or norms about education, determined by randomization. The questions reported in columns 1 and 4, which ask about personal attitudes, are not included in the gender attitudes index. Standard errors are clustered by school.

change, we use them to examine the extent to which participants view social norms as preventing them from acting on their progressive attitudes.

Among girls, the intervention made personal attitudes about female employment more progressive by 8 percentage points (Table 6, column 1) but did not significantly increase their perception that others in the community hold that gender-progressive view (column 2). In contrast, among boys, not only is there a treatment effect on their personal attitude, but there is also a significant increase in how progressive they view the community to be. Column 3 shows that, among both girls and boys, the treatment group is more likely to hold a progressive attitude and believe society will be supportive; many of those whose gender attitudes became more progressive do not expect the community norms to oppose their views. However, the estimates in column 3 are appreciably smaller than the estimates for personally holding that attitude reported in column 1. Some students whose own attitude changed think that restrictive cultural norms will hinder them from acting on it. Columns 4 to 6 show a broadly similar pattern regarding the social norm about women leaving home to attend college, with the intervention only changing boys' perceptions of social norms.

Another secondary outcome is school performance, which we examine to rule out the concern that the program hurt achievement in subjects like math and Hindi by taking away some of their instructional time. We find no impact on school performance, as reported in online Appendix Table 20. The administrative exam-score data we use are aggregated at the school-grade level, so we cannot estimate results separately by gender. However, on average, girls outperform boys on secondary

TABLE 7—TREATMENT EFFECTS ON OTHER SECONDARY OUTCOMES (END LINE 1)

	Girls' self-esteem (1)	Awareness of gender-based discrimination (2)	IAT: associates girls with positive words (3)	IAT: associates women with market work (4)
Panel A. Girls				
Treated	0.104	0.053	−0.006	−0.079
	[0.023]	[0.021]	[0.047]	[0.074]
Control group mean	0.000	0.099	0.408	0.000
Number of students	7,788	7,777	1,676	1,830
Panel B. Boys				
Treated	N/A	0.007	0.014	−0.004
	N/A	[0.020]	[0.048]	[0.063]
Control group mean	N/A	−0.118	−0.514	−0.000
Number of students	N/A	6,162	1,250	1,368

Notes: All regressions control for grade and district fixed effects. All columns except column 2 also control for the baseline analog of the outcome. Columns 1 and 2 also control for missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

school exams in Haryana. This fact combined with girls' high aspirations to attend college meant that improvements in girls' school performance was not an expected impact of the program.

Results for the three other prespecified secondary outcomes (described in more detail in the online data Appendix) are reported in Table 7. First, we find that the program improved girls' self-esteem (Rosenberg 1965). Second, we find a small increase in awareness of gender discrimination among girls but not boys; status-quo awareness is already high. Third, we find no impact on two implicit association tests (IATs), each of which was administered to about 3,000 respondents.[21] One IAT measured how respondents associate girls' and boys' faces with positive and negative words. The other, originally developed by Beaman et al. (2009), associates men's and women's faces with market labor or domestic work. Because of challenges administering the IATs at baseline (13 percent of responses were invalidated because the completion time was too fast or slow), we did not make the IAT a key focus for the end line.[22]

V. Medium-Run Results

We next investigate whether the effects described above persisted. These results use data from the second end line survey, which was conducted 2 to 2.5 years after the program ended.

---

[21] At baseline, we administered an IAT to a randomly selected 50 percent of the sample, and in the first end line we readministered an IAT to this subsample, with half receiving each version of the IAT. Some end line IAT observations are missing due to technical problems with the laptops used.

[22] Another concern with IATs is that they measure not only personal attitudes but also awareness of shared cultural stereotypes. Thus, a program that discusses stereotypes could lead to a "worse" IAT score (Arkes and Tetlock 2004).

TABLE 8—TREATMENT EFFECTS ON ATTITUDES, ASPIRATIONS, AND BEHAVIOR (END LINE 2)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)	Applied to scholarship (4)	Signed petition (5)
Treated	0.160 [0.019]	−0.025 [0.019]	0.227 [0.025]	0.031 [0.017]	0.012 [0.009]
Control group mean	0.333	0.000	0.000	0.408	0.150
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	13,679	7,560	13,677	7,347	13,303

Notes: All regressions control for grade-gender and district-gender fixed effects (columns 1, 3, and 5) or grade and district fixed effects (columns 2 and 4). Columns 1 to 3 also control for the baseline analog of the outcome and missing flags for each variable used to construct the outcome index. The outcomes in columns 4 and 5 are binary variables, not indices, and were not collected for the 3 percent of respondents who were surveyed by phone for the second end line (because these outcomes involved giving printed material to the respondent). Standard errors are clustered by school.

A. *Effects on Primary Outcomes, Including Heterogeneity by Gender*

We continue to find a large and significant effect on gender attitudes in the second end line, as reported in column 1 of Table 8. The attitudes index is constructed identically to the first end line's index, so the 0.16 effect size medium-run effect is directly comparable to the 0.18 short-run effect size. The estimate is robust to including double-LASSO-selected controls (online Appendix Table 21) and using Lee bounds (online Appendix Table 10). The control group mean of 0.33 in the second end line indicates that attitudes became more progressive between the end lines, absent the intervention. Thus, the 11 percent fade-out in the treatment effect is not because the treatment group held less progressive attitudes at the second end line than at the first end line; rather, their attitudes improved less between the two waves than the control group's attitudes did.

Experimenter demand effects are somewhat less of a concern in the second end line because the intervention had ended two years earlier and so was less likely to be top of mind for the treatment group. Nonetheless, it is important to investigate this potential confound and we do so using the same approach as earlier. We find that having a high propensity to give socially desirable answers is not associated with having larger treatment effects. That is, the coefficient on the interaction of *Treated* and having a high social desirability score is small and statistically insignificant, as shown in Table 9.

Turning to the results by gender, while there was no significant difference in attitude change between boys and girls in the short run, two years later there is. The medium-run effect size on boys' gender attitudes is 0.22, as shown in Table 10, column 2. This point estimate is slightly larger than their short-run effect size: there is no fade-out for boys. The effect for girls is 0.11 standard deviations and statistically significant ($p < 0.01$). This is two-thirds of the short-run effect size for girls, though we cannot statistically rule out identical effect sizes across the two end lines. One conjecture is that acting on one's beliefs reinforces them, such that the gender gap in the persistence of attitude change is related to the smaller change in behavior among girls that we observed in the short run. In any case, we view this pattern as interesting and worthy of future research.

Table 9—Robustness Check for Social Desirability Bias (End Line 2)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)	Applied to scholarship (4)	Signed petition (5)
Treated	0.150 [0.024]	−0.034 [0.024]	0.235 [0.028]	0.034 [0.019]	0.020 [0.010]
High social desirability score	0.070 [0.022]	0.029 [0.025]	0.059 [0.024]	0.017 [0.017]	0.014 [0.008]
Treated × High Soc. D score	0.028 [0.031]	0.024 [0.034]	−0.021 [0.034]	−0.006 [0.025]	−0.021 [0.013]
p-value: Treated + Treated × High Soc. D = 0	0.000	0.728	0.000	0.230	0.946
Control group mean	0.333	0.000	0.000	0.406	0.150
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	13,679	7,560	13,677	7,347	13,303

Notes: Social desirability (Soc. D) score is a baseline measure of the student's propensity to give socially desirable answers. High Soc. D score refers to having an above-median score among students. All regressions control for grade-gender and district-gender fixed effects (columns 1, 3, and 5) or grade and district fixed effects (columns 2 and 4). Columns 1 to 3 also control for the baseline analog of the outcome and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 10—Gender-Specific Treatment Effects on Attitudes, Aspirations, and Behavior (End Line 2)

	Gender attitudes index		Self-reported behavior index		Signed petition	
	Girls (1)	Boys (2)	Girls (3)	Boys (4)	Girls (5)	Boys (6)
Treated	0.111 [0.025]	0.218 [0.028]	0.158 [0.025]	0.311 [0.040]	0.019 [0.013]	0.003 [0.010]
Control group mean	0.562	0.063	−0.067	0.079	0.189	0.104
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of students	7,562	6,117	7,563	6,114	7,347	5,956
p-value: Girls = Boys	0.003		0.000		0.302	

Notes: All regressions control for grade and district fixed effects, the baseline analog of the outcome, and missing flags for each variable used to construct the outcome index. The outcome in columns 5 and 6 is a binary variable, not an index, and was not collected for the 3 percent of respondents who were surveyed by phone for the second end line (because these outcomes involved giving printed material to the respondent). Standard errors are clustered by school.

A treatment effect on girls' aspirations could have emerged between the two end lines if, in the control group, girls lowered their aspirations over time. This is not the case: as in the short run, we find no effect of the intervention on girls' stated aspirations in the medium run (Table 8, column 2).

For self-reported behavior, we find a sizable (0.23 standard deviations) and statistically significant treatment effect, pooled for boys and girls, just as we did in the first end line. Note that we updated the elements in the behavior index between the two end lines, so the effect size is not as directly comparable over time as the effect on attitudes is. The treatment effect on self-reported behavior continues to be larger for boys than girls in the medium run (Table 10, columns 3 and 4). Our interpretation of this pattern in the short run was that girls were more constrained by external factors in translating attitudes to behavior. Now another contributing factor is that attitude change is less sustained for girls.

We next examine treatment effects on the two revealed-preference outcomes that we introduced in the second end line. The first measure, for girls only, is applying for a college scholarship. The theory of change is that the program either made girls' desire to attend college more intense (higher aspirations) or enabled them to persuade their parents to support their goal (changed expectations about their future behavior, conditional on aspirations), making it more worth their while to complete the application. We find that the intervention led to a marginally significant increase of 3.1 percentage points, or 8 percent, in the application rate ($p = 0.07$), as reported in Table 8, column 4). However, the p-value for this treatment effect increases to 0.13 in the alternative specification using double-LASSO-selected control variables (see online Appendix Table 21).

The intervention could have affected scholarship applications either by strengthening girls' resolve to go to college or by enabling them to secure their parents' support. That is, it could have changed their preferences or relaxed a constraint. We conduct some exploratory heterogeneity analysis to further probe this. First, we find that the treatment effect on scholarship take-up is significantly higher for girls who had higher aspirations at baseline, as reported in Table 11, column 1. The total effect is close to four times as high among those with above-median initial aspirations compared to those with below-median aspirations (column 2). Second, the treatment effect on scholarship applications is significantly higher among the 80 percent of girls who, at baseline, said they had talked with their parents about their education goals (column 3). If the results had been concentrated among those with low aspirations, it would have been suggestive that girls' resolve to attend college increased. If it had been concentrated among those with low parental engagement, this would have pointed to girls securing their parents' support. Instead, these results do not clearly point to just one of these mechanisms operating. The intervention seems to have increased girls' intention to attend college by raising already-high aspirations and convincing already-engaged parents to support their daughter's goals, rather than by converting girls and families who started out lower on these dimensions.

Our final primary outcome is signing a public petition to end the dowry system. The intervention might have either made participants more opposed to the dowry system (attitude change) or reduced the cost to them of expressing their attitude (change in perceived social sanctions). In the control group, 15 percent of respondents called to add their names to the petition.[23] This rate is not significantly different in the treatment group, though the point estimate is in the direction of a small increase, as shown in Table 8, column 5. This positive point estimate is driven mostly by girls (see Table 10, columns 5 and 6). In the alternative specification using double-LASSO-selected controls, the treatment effect on signing the petition is marginally significant, with $p = 0.07$ (see online Appendix Table 21). Thus, we find weak evidence that the intervention influenced this outcome. One way to reconcile this weak evidence with the strong observed effect on self-reported opposition to the dowry system is that the self-reported attitude change is disingenuous.

---

[23] Ideally, for statistical power, the measure would have had a higher mean in the control group; the estimated effect size on self-reported attitudes corresponds to a large (38 percent) increase in the likelihood of signing. The requirement of needing to phone in may have been an impediment to signing. We chose not to have respondents sign in the presence of the surveyor, as that variant could still be susceptible to experimenter demand effects, which would defeat the purpose of adding this outcome.

TABLE 11—UNPACKING THE TREATMENT EFFECT ON SCHOLARSHIP APPLICATIONS (END LINE 2)

| | Applied to scholarship | | |
	(1)	(2)	(3)
Treated	0.029	0.014	−0.023
	[0.017]	[0.019]	[0.027]
Treated × BL aspirations index	0.022		
	[0.011]		
Treated × Above-median BL aspirations		0.040	
		[0.024]	
Treated × Has discussed educ goals with parent			0.068
			[0.028]
p-value: Treated + Treated × Above-median aspirations = 0		0.020	
p-value: Treated + Treated × Has discussed goals = 0			0.016
Control group mean	0.408	0.408	0.408
Number of students	7,347	7,347	7,347

Notes: All regressions include grade and district fixed effects, the main effects for the baseline variable used in the interaction term, and flags for whether the baseline variable is missing. Standard errors are clustered by school.

Another possibility is that students worried about repercussions in their family or community from a public statement of their view, and the intervention did not lower this perceived cost (enough). The analysis using the Marlowe-Crowne measure (Table 9, column 5) shows an interesting pattern for this outcome. The intervention significantly increased petition signing by 2 percentage points in the subsample with low social desirability concerns, as seen from the main effect of *Treated*. In addition, in the control group but not the treatment group, those with strong social desirability concerns were more likely to sign the petition; one reason for signing the petition may have been to look good to others.[24] These results are consistent with there being a set of people for whom social desirability concerns and genuine opposition to the dowry system instilled by the intervention were substitutes, with either being sufficient to sign the petition.[25]

B. *Effects on Secondary Outcomes*

We reexamine perceived social norms as a secondary outcome in the second end line, with the results reported in Table 12. We continue to find that among boys but not girls, the treatment group views the community as more supportive of female employment, but there is no longer an effect on perceived norms about female education.

In addition, we continue to see a positive effect on girls' self-esteem. This result is reported in Table 13, along with results for the remaining secondary outcomes. The third secondary outcome is an index of girls' education outcomes. It is based

[24] In online Appendix Table 20 we test whether the stated and revealed preference measures are less correlated with each other among those with a high social desirability score.

[25] In the online Appendix we present some further medium-run analyses. Online Appendix Table 22 assesses social desirability bias separately by gender. Online Appendix Tables 23 to 26 show the results for thematic sub-indices and the individual variables that comprise the indices. Based on the lack of robust heterogeneous results by parental attitudes at the first end line, we did not prespecify it as a dimension of heterogeneity we would analyze for the second end line. Nonetheless, for completeness, we report these (null) results in online Appendix Table 27.

TABLE 12—TREATMENT EFFECTS ON PERCEPTIONS OF SOCIAL NORMS (END LINE 2)

| | Social norms toward work | | | Social norms toward education | | |
| | *Student agrees:* | | | *Student agrees:* | | |
	Women should be allowed to work (1)	Community thinks women should be allowed to work (2)	Women should be allowed to work and thinks community will not oppose them (3)	Women should be allowed to study in college even if it is far away (4)	Community thinks women should be allowed to study in college even if it is far away (5)	Women should be allowed to study in college and thinks community will not oppose them (6)
Panel A. Girls						
Treated	0.013	0.005	0.006	0.011	−0.009	−0.011
	[0.006]	[0.019]	[0.018]	[0.008]	[0.019]	[0.017]
Control group mean	0.965	0.643	0.707	0.950	0.649	0.712
Number of students	3,590	3,435	3,418	3,542	3,403	3,378
Panel B. Boys						
Treated	0.119	0.070	0.092	0.051	0.027	0.038
	[0.016]	[0.019]	[0.019]	[0.012]	[0.017]	[0.017]
Control group mean	0.747	0.576	0.577	0.866	0.708	0.719
Number of students	3,043	2,945	2,935	2,902	2,808	2,801
Panel C. Girls = boys p-value						
	0.000	0.007	0.000	0.006	0.148	0.048

Notes: All columns control for grade and district fixed effects. Standard errors are clustered by school.

on school enrollment, studying a STEM subject, taking classes to obtain extra skills (e.g., computer, English), and taking after-school tutoring for exam preparation. We see a marginally significant increase of 0.06 standard deviations in the index (Table 13, column 2).

We also added a set of questions on aspirations related to age of marriage and son preference. For both boys and girls, we find a modest increase in an index of these marriage and fertility aspirations, in the direction of more gender progressiveness (columns 3 and 4).

The final two secondary outcomes are related to sexual harassment. We asked girls about the harassment they experienced, and we asked boys if they engaged in sexual harassment and assault. Because of concern about underreporting by boys, we used an item count (or list) experiment for them, with one group receiving an extra item that asked them if the following statement is true: "In the past year, I have passed dirty comments about a girl; made dirty gestures in a girl's presence, or inappropriately touched or groped a girl." We find that the intervention increased girls' reported harassment (column 5), and though we do not find a significant effect on boys' reported engagement in harassment, the point estimate is positive (column 6). While we cannot disentangle whether these patterns are due to actual experiences or to changes in reporting, we speculate that the intervention made both boys and girls more aware of harassment, so they recognized and reported harassment at a higher rate. An important area for further work is to understand if these patterns, instead, reflect actual increases in harassment, which might arise from the greater interaction among boys and girls that the program encouraged.

Table 13—Treatment Effects on Other Secondary Outcomes (End Line 2)

	Girls' self-esteem (1)	Girls' education (2)	Marriage and fertility aspirations (girls) (3)	Marriage and fertility aspirations (boys) (4)	Girls' experienced harassment (5)	Boys' perpetrated harassment (school-grade level) (6)
Treated	0.086	0.058	0.052	0.047	0.063	0.060
	[0.026]	[0.033]	[0.029]	[0.028]	[0.029]	[0.062]
Control group mean	0.000	0.000	0.143	0.169	0.000	−0.003
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,341	7,566	7,369	5,919	7,314	504

Notes: The unit of observation is the student in columns 1 to 5 and the school-grade in column 6. All columns control for grade and district fixed effects and, for columns 1 to 4, missing flags for each variable used to construct the outcome index. Column 1 also controls for the baseline analog of the outcome variable. Standard errors are clustered by school. A higher value of the outcome in column 5 corresponds to higher reported harassment by girls. The outcome in column 6 is the proportion of boys who report engaging in sexual harassment, based on a list experiment.

VI. Conclusion

This paper studies an attempt to change people's views about how acceptable gender inequality is and to erode their support for societal norms that restrict women's and girls' opportunities. We examined whether an attitude-change intervention in schools in Haryana, India, could make adolescents' attitudes and, in turn, their behavior, less gender biased. The intervention had students discuss and think about gender norms and gender discrimination through class sessions held over two and one-half school years.

We find that the intervention succeeded in making attitudes more supportive of gender equality and coaxing more gender-equal behavior. Moreover, the impacts persisted: When we resurveyed participants two years after the program had ended, we continued to find strong effects.

North India has particularly strong gender discrimination, and gender norms are often highly dependent on the context. Thus, the exact programming would need to be adapted for other contexts, but this general approach of engaging adolescents, or even younger children, in school discussions could be a widely applicable way to change gender norms. While it might seem surprising that a series of class discussions changed views formed over many years, most adolescents had likely spent very little time thinking hard about gender inequality prior to the program. Interrogating one's personal prejudices and the norms that prevail in one's community might be especially powerful in schools, without voluntary, self-selected participation and when participants are young enough that their views are still quite pliable.

That said, perhaps a more indirect approach would be even more effective than explicit discussions about the importance of gender equality; assigned readings could feature empowered women, or a history assignment could ask students to discuss women's changing role in society. This aspect of the program as well as other ones, such as participants' age, the dosage, and the relative emphasis on moral versus pragmatic arguments for equality, could be varied and assessed to optimize programs like this one.

One lesson from our results is the importance of including boys and men in programs aimed at altering gender norms. We find that attitude change translates into larger shifts in behavior for boys than girls. We speculate that the very problem the program aims to solve—that males have more power in society—means that they have more freedom to act on the gender-progressive views the program instilled in them. For this reason, as we look ahead, we expect a larger increase in employment for male participants' wives than for female participants. Of course, men also face familial constraints on their behavior and feel pressure to conform to traditional norms, so it is an open question whether such an effect on wives' employment will materialize. In future work, we hope to measure whether it indeed does, as well as how the program affects other adult outcomes such as higher education, age of marriage, and childbearing.

REFERENCES

Abramsky, Tanya, Karen Devries, Ligia Kiss, Janet Nakuti, Nambusi Kyegombe, Elizabeth Starmann, Bonnie Cundill et al. 2014. "Findings from the SASA! Study: A Cluster Randomized Controlled Trial to Assess the Impact of a Community Mobilization Intervention to Prevent Violence Against Women and Reduce HIV Risk in Kampala, Uganda." *BMC Medicine* 12 (1): 122.

Afridi, Farzana, Taryn Dinkelman, and Kanika Mahajan. 2018. "Why Are Fewer Married Women Working in Rural India? A Decomposition Analysis Over Two Decades." *Journal of Population Economics* 31 (3): 783–818.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics* 128 (2): 469–530.

Aloud, Monira Essa, Sara Al-Rashood, Ina Ganguli, and Basit Zafar. 2020. "Information and Social Norms: Experimental Evidence on the Labor Market Aspirations of Saudi Women." NBER Working Paper 26693.

Alrababa'h, Ala', William Marble, Salma Mousa, and Alexandra Siegel. 2019. "Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes." Immigration Policy Lab Working Paper 19-04.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Arkes, Hal R., and Philip E. Tetlock. 2004. "Attributions of Implicit Prejudice, or "Would Jesse Jackson 'Fail' the Implicit Association Test?" *Psychological Inquiry* 15 (4): 257–78.

Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn. 2020. "Negotiating a Better Future: How Interpersonal Skills Facilitate Inter-Generational Investment." *Quarterly Journal of Economics* 135 (2): 1095–151.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2020. "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa." *American Economic Journal: Applied Economics* 12 (1): 210–59.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of Economics* 124 (4): 1497–1540.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection Among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–50.

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer-Moore. 2018. "Male Social Status and Women's Work." *AEA Papers and Proceedings* 108: 363–67.

Bicchieri, Cristina. 2016. *Norms in the Wild: How to Diagnose, Measure and Change Social Norms.* Oxford: Oxford University Press.

Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, and Iman Sen. 2018. "Power vs. Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial." Unpublished.

Bursztyn, Leonardo, Alessandra L. Gonzalez, and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110 (10): 2997–3029.

Cantoni, Davide, Yuyu Chen, David Yang, Noam Yuchtman, and Y. Zhang. 2017. "Curriculum and Ideology." *Journal of Political Economy* 125 (2): 338–92.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *Quarterly Journal of Economics* 134 (3): 1163–1224.

Carrell, Scott E., Mark Hoekstra, and James E. West. 2019. "The Impact of College Diversity on Behavior Toward Minorities." *American Economic Journal: Economic Policy* 11 (4): 159–82.

Crowne, Douglas P., and David Marlowe. 1960. "A New Scale of Social Desirability Independent of Psychopathology." *Journal of Consulting Psychology* 24 (4): 349–54.

Dahl, Gordon B., Andreas Kotsadam, and Dan-Olof Rooth. 2021. "Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams." *Quarterly Journal of Economics* 136 (2): 987–1030.

DellaVigna, Stefano, and Matthew Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics* 2 (1): 643–69.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2014. "Evaluation of a School-Based Gender Sensitization Campaign in India." AEA RCT Registry. February 26. https://doi.org/10.1257/rct.72-11.0.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2019. "The Intergenerational Transmission of Gender Attitudes: Evidence from India." *Journal of Development Studies* 55 (12): 2572–92.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2022. "Replication data for: Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." American Economic Association [Publisher] Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E149882V1.

DISE. 2011. "District Information System for Education." Unified District Information System for Education. http://dise.in (accessed September 19, 2018).

Donovan, Robert J., and Susan Leivers. 1993. "Using Paid Advertising to Modify Racial Stereotype Beliefs." *Public Opinion Quarterly* 57 (2): 205–18.

Duflo, Esther. 2012. "Women Empowerment and Economic Development." *Journal of Economic Literature* 50 (4): 1051–79.

Edmonds, Eric V., Benjamin Feigenberg, and Jessica Leight. 2020. "Advancing the Agency of Adolescent Girls." Unpublished.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics* 119 (4): 1249–99.

Field, Erica, Seema Jayachandran, and Rohini Pande. 2010. "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *American Economic Review Papers and Proceedings* 100 (2): 125–29.

Government of India. 2011. "Census of India." Office of the Registrar General and Census Commissioner, India. www.censusindia.gov.in/.

Green, Donald P., Anna M. Wilke, and Jasper Cooper. 2020. "Countering Violence Against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda." *Comparative Political Studies* 53 (14): 2283–320.

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal. 2020. "Does Information Change Attitudes Toward Immigrants?" *Demography* 57 (3): 1117–43.

Gupta, Jhumka, Kathryn Falb, Heidi Lehmann, Denise Kpebo, Ziming Xuan, Mazeda Hossain, Cathy Zimmerman, Charlotte Watts, and Jeannie Annan. 2013. "Gender Norms and Economic Empowerment Intervention to Reduce Intimate Partner Violence Against Women in Rural Côte d'Ivoire: A Randomized Controlled Pilot Study." *BMC International Health and Human Rights* 13 (46).

Healy, Andrew, and Neil Malhotra. 2013. "Childhood Socialization and Political Attitudes: Evidence from a Natural Experiment." *Journal of Politics* 75 (4): 1023–37.

Hopkins, Daniel J., John Sides, and Jack Citrin. 2019. "The Muted Consequences of Correct Information about Immigration." *Journal of Politics* 81 (1): 315–20.

Jayachandran, Seema. 2015. "The Root of Gender Inequality in Developing Countries." *Annual Review of Economics* 7 (1): 63–88.

Jayachandran, Seema. 2021. "Social Norms as a Barrier to Women's Employment in Developing Countries." *IMF Economic Review* 69 (3): 576–95.

Jayachandran, Seema. 2017. "Fertility Decline and Missing Women." *American Economic Journal: Applied Economics* 9 (1): 118–39.

Jensen, Robert, and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *Quarterly Journal of Economics* 124 (3): 1057–94.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin. 2006. "Low Male-to-Female Sex Ratio of Children Born in India: National Survey of 1.1 million Households." *Lancet* 367 (9506): 211–18.

Kishor, Sunita, and Kamla Gupta. 2004. "Women's Empowerment in India and its States: Evidence from the NFHS." *Economic and Political Weekly* 39 (7): 694–712.

Klasen, Stephan, and Janneke Pieters. 2015. "What Explains the Stagnation of Female Labor Force Participation in Urban India?" *World Bank Economic Review* 29 (3): 449–78.

Kohlberg, Lawrence. 1976. "Moral Stages and Moralization: The Cognitive-Developmental Approach." In *Moral Development and Behavior: Theory, Research, and Social Issues*, edited by Thomas Lickona, 31–53. New York: Holt, Rinehart, and Winston.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4 (4): 1–31.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76 (3): 1071–1102.

Lowe, Matt. 2021. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *American Economic Review* 111 (6): 1807–44.

Markus, Hazel, and Paula Nurius. 1986. "Possible Selves." *American Psychologist* 41 (9): 954–969.

McKelway, Madeline. 2020. "Women's Employment in India: Intra-Household and Intra-Personal Constraints." Unpublished.

Menon, Sunita. 2021. "Curriculum for a Gender Equal World." *Hindustan Times*, September 5.

Moss-Racusin, Corrine, Evava Pietri, Erin Hennes, John Dovidio, Victoria Brescoll, Gina Roussos, and Jo Handelsman. 2018. "Reducing STEM Gender Bias with VIDS (Video Interventions for Diversity In STEM)." *Journal of Experimental Psychology: Applied* 24 (2): 236–60.

Mukherjee, Bishwa Nath. 1967. "A Cross-Validation of the Marlowe-Crowne Social Desirability Scale on an Indian Sample." *Journal of Social Psychology* 72 (2): 299–300.

Paluck, Elizabeth. 2009. "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda." *Journal of Personality and Social Psychology* 96 (3): 574–87.

Porter, Catherine, and Danila Serra. 2020. "Gender Differences in the Choice of Major: The Importance of Female Role Models." *American Economic Journal: Applied Economics* 12 (3): 226–54.

Pulerwitz, Julie, Lindsay Hughes, Manisha Mehta, Aklilu Kidanu, Fabio Verani, and Samuel Tewolde. 2015. "Changing Gender Norms and Reducing Intimate Partner Violence: Results from a Quasi-Experimental Intervention Study with Young Men in Ethiopia." *American Journal of Public Health* 105 (1): 132–37.

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

Reynolds, William. 1982. "Development of Reliable and Valid Short Forms of the Marlowe-Crowne Social Desirability Scale." *Journal of Clinical Psychology* 38 (1): 119–25.

Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Sen, Amartya. 1990. "More Than 100 million Women Are Missing." *New York Review of Books.* December 20.

Tankard, Margaret, and Elizabeth Paluck. 2016. "Norm Perception as a Vehicle for Social Change." *Social Issues and Policy Review* 10 (1): 181–211.

Vu, Alexander, Nhan Tran, Kiemanh Pham, and Saifuddin Ahmed. 2011. "Reliability of the Marlowe-Crowne Social Desirability Scale in Ethiopia, Kenya, Mozambique, and Uganda." *BMC Medical Research Methodology* 11 (1): 162.

Washington, Ebonya. 2008. "Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues." *American Economic Review* 98 (1): 311–32.

World Bank. 2011. "World Development Indicators." World Bank.

This article has been cited by:

1. Sonali Rakshit, Soham Sahoo. 2023. Biased teachers and gender gap in learning outcomes: Evidence from India. *Journal of Development Economics* **161**, 103041. [Crossref]

2. Can Tang, Liqiu Zhao. 2023. Gender Social Norms and Gender Gap in Math: Evidence and Mechanisms. *Applied Economics* 1-19. [Crossref]

3. Shelly Lundberg. Gender Economics: Dead-Ends and New Opportunities 151-189. [Crossref]

4. Saravana Ravindran, Manisha Shah. 2023. Unintended consequences of lockdowns, COVID-19 and the Shadow Pandemic in India. *Nature Human Behaviour* **340**. . [Crossref]

5. Upasak Das, Karan Singhal. 2023. Solving it correctly: Prevalence and persistence of gender gap in basic mathematics in rural India. *International Journal of Educational Development* **96**, 102703. [Crossref]

6. Marisol Rodríguez Chatruc, Sandra Viviana Rozo. 2022. Discrimination toward migrants during crises. *Migration Studies* **10**:4, 582-607. [Crossref]

7. Siwan Anderson. 2022. Unbundling female empowerment. *Canadian Journal of Economics/Revue canadienne d'économique* **55**:4, 1671-1701. [Crossref]

8. Dante Donati, Victor Orozco-Olvera, Nandan Rao. Using Social Media to Change Gender Norms: An Experiment within Facebook Messenger in India **6**, . [Crossref]

9. Tarun Jain, Diva Dhar, Vrinda Kapoor, Vrinda Kapur, Anita Raj. 2022. Measuring gender attitudes: Developing and testing Implicit Association Tests for adolescents in India. *PLOS ONE* **17**:6, e0264077. [Crossref]

10. Marisol Rodriguez Chatruc, Sandra V. Rozo. Discrimination toward Migrants during Crises **103**, . [Crossref]