

Transmitting Rights: Effective Cooperation, Inter-Gender Contact, and Student Achievement

By Sultan Mehmood, Shaheen Naseer and Daniel L. Chen¹

July 2024

We provide experimental evidence of teacher-to-student transmission of gender attitudes in Pakistan. We randomly show teachers a pro-women’s rights visual narrative. Treated teachers increase their and students’ support for women’s rights, unbiasedness in gender IATs, and willingness to petition parliament for greater gender equality. Students improve coordination and cooperation with the opposite gender. Effects are larger when teachers teach a gender-rights curriculum. Mathematics achievement increases for classrooms assigned to form mixed-gender study groups treated with an intense program (visual narrative and curriculum), while no significant effects appear in same-sex study groups. Gender attitudes are transmissible, and cooperation improves student outcomes.

JEL Codes: I24, I28, O12, C93

Keywords: teachers, attitudes, IATs, gender, inter-gender contact.

¹ New Economic School (E-mail: smehmood@nes.ru). Lahore School of Economics (E-mail:ranashaheen133@gmail.com) and Toulouse School of Economics (Email: daniel.chen@iast.fr). Daniel L. Chen acknowledges IAST funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d’Avenir) program, grant ANR-17-EUR-0010. This research received financial support from the research foundation TSE- Partnership and ANITI funding. We would like to thank Director Training, Miss Sumera Morris, and Director Academics and Curriculum Development Miss Faria Hafeez at the Progressive Education Network for their invaluable support and cooperation throughout the experiment. We also thank Avner Seror, Christina Brown, Seema Jayachandran, Peter Hull, Michela Carlana, Marjorie Sweetko, Alberto Bisin, Matthias Doepke, Gautam Rao, Roberta Ziparo, Hosny Zoabi, Abhijit Banerjee, Esther Duflo, Thierry Verdier, and Eliana La Ferrara for their comments and feedback. Sameen Tariq, Sajwaar Khalid and Bakhawar Ali provided excellent research assistance and on-the-field support. Pre-registered with AEA RCT Registry with ID: AEARCTR-0007465. The IRB obtained from the Lahore School of Economics with IRB Number RERC-062021-03.

"A teacher affects eternity." Henry Adams in A Letter to Teachers (1910)

I. Introduction

How are rights revolutions transmitted in societies? The last century witnessed an extraordinary extension of civil rights and freedoms to women and to religious, sexual, and ethnic minorities. Particularly impressive progress was achieved on women's rights, especially in economically developed societies. Yet in much of the world, women still have fewer labor market and educational opportunities, lower physical mobility, less autonomy to run for political office or to make their own decisions about marriage, divorce, finances, and even less freedom to choose their friends (Doepke and Tertilt, 2009; Duflo, 2012; Fernández, 2014; Fernandez and Wong, 2014; Giuliano, 2020; Field et al., 2021; Seror, 2022). These disparities are often rooted in cultural attitudes and transmitted from generation to generation (Maurin and McNally, 2008; Bisin and Verdier, 2011; Alesina et al., 2013; Doepke and Zilibotti, 2017; Giuliano and Nunn, 2021).

Given these entrenched disparities, our study advances the discourse on gender parity by empirically substantiating the transformative power of education in reforming societal norms. We corroborate findings from influential randomized control trials (Alan et al., 2021; Dhar et al., 2022), which demonstrate that curricular reforms can significantly foster equitable gender attitudes and enhance social cohesion. Building on this literature, our study offers three core insights into the effectiveness of such interventions. First, non-gender-specific interventions, such as those aimed at fostering overall empathy, fail to significantly alter gender attitudes or improve student test scores, suggesting the ineffectiveness of broad-based strategies. Second, gender-focused interventions, even light-touch ones like a gender-themed film screening with structured discussion, can positively influence gender attitudes, suggesting the efficacy of

targeted approaches for influencing attitudes. Third, to significantly influence attitudes and improve academic outcomes, intensive measures such as pairing a gender-rights movie with a related curriculum are needed to foster inter-gender cooperation in education and enhance students' test scores.

We implement a randomized control trial in collaboration with the Progressive Education Network (PEN), one of the largest networks of charter schools in the world, operating schools across Pakistan. We randomly assign female teachers to one of the following four treatment arms. The first two treatment arms are inspired by economic theory: utilitarianism and identity theory. The principle of utilitarianism is that actions should be evaluated by their utility for oneself and for society as a whole. By emphasizing the importance of empathy, teachers may be encouraged to embrace a more balanced approach to gender equality. Teaching the malleability of one's identity, in particular regarding growth in empathy, may increase empathy toward outgroups (see e.g., Weisz and Zaki, 2017). In the third treatment, teachers are presented with a visual narrative (Riley, 2019; Banerjee et al., 2019) developed at Johns Hopkins University—an award-winning movie—emphasizing the importance of women's rights (Benabou, Falk, and Tirole, 2018). Finally, the fourth treatment reinforces the visual narrative with a gender-rights curriculum that the teachers then teach to their students in a semester-long class.

We measure the impacts both on the teacher's own attitudes and on their students' attitudes, as well as on student achievement and gender gaps in student achievement. The gender-rights lessons included structured discussions where students envisioned the rights of men and women: the curriculum encouraged self-reflection by students. To create the opportunity for reflective equilibrium—a deliberative mutual adjustment through inter-gender interactions—we cross-randomized teachers to form either mixed-gender or same-gender student study groups. This cross-randomization enabled us to experimentally examine a

causal mediating channel (of inter-gender contact) through which achievement may be boosted. Outcomes are measured six months and one year after the treatment.

The visual narrative presented in the 'Bol' movie influences gender rights attitudes among teachers in the short and medium term. Treatment effects are observed after both the stand-alone visual narrative and the joint visual narrative and gender-rights curriculum. Support for equitable gender rights among teachers was roughly 0.15 standard deviations higher in the group treated with the visual narrative *alone* and increases to 0.25 standard deviations for teachers when the visual narrative is combined with the gender-rights curriculum. Effects were persistent, with coefficient estimates remaining similar about six months and one year after the treatment. The shift in attitudes is also reflected in the teachers' decisions to petition the government. The visual narrative treatment alone led to at least a 5 percentage point increase in the likelihood of signing pro-women's rights petitions. When this narrative was paired with the gender-rights curriculum, the coefficient estimates suggest that the likelihood of petitioning rose by approximately 20 percentage points. This is at least doubling of petitions over the sample mean relative to the control group. Teachers' implicit attitudes, as measured by IAT scores, are also affected: the joint treatment reduced implicit gender bias by about 0.35 standard deviations in both the short and medium term.

Consistent with teachers' internalizing the shift in attitudes through their teaching, the stand-alone visual narrative treatment spilled over to students, and impacted their gender attitudes. The visual narrative treatment on its own led to students becoming 0.1 standard deviations more supportive of gender rights. When it was combined with the semester-long gender-rights curriculum, students became more than twice as likely —about 0.25 standard deviations— more supportive of gender rights. These two results together illustrate the teacher-to-students transmission of gender rights attitudes, especially striking is

the impact of the standalone visual narrative treatment on students since this treatment *squarely* focused on teachers. Students' behavior in cooperation and coordination games when the interaction partner is the opposite gender shifted by 0.2 (in the visual narrative treatment arm) to 0.3 standard deviations (in the visual narrative plus gender-rights curriculum treatment arm). Students in the joint treatment arm (visual narrative plus gender-rights curriculum) also score 0.12 standard deviation higher in mathematics exams. All of these effects, the elevation in cooperation and coordination to the opposite gender due to visual narrative treatment or the joint treatment as well as the elevation in math achievement due to the joint treatment, arise only in mixed-gender study groups.

The rest of the paper is organized as follows. Section II provides the background and experimental design. Section III describes the data and empirical specification, while Sections IV and V present the main results. Section VI reports the results of our experimental mediation analysis. A final section provides some concluding remarks. Online Appendices discusses a series of sensitivity tests, provides the corresponding results and additional experimental details.

II. Background and Study Design

Background.—We collaborate with the Progressive Education Network (PEN) and embed a large field experiment within their *Teacher Training Drives 2021*. PEN is a Non-Profit-Organization (NGO) that aims to improve the quality of education via a public-private “charter”. The approach is similar to charter schools in the United States. These schools are privately managed using public funds, in a public-private partnership. We implement a randomized evaluation in all charter schools in Punjab, the largest province of Pakistan, where the network employs 607 teachers and oversees about 15000 students. According to the PEN mission statement, the focus is not on infrastructure investments but rather on raising the quality of education by improving the quality of teachers:

“We believe that instead of investing resources in ‘bricks and mortar’, we can leverage the existing infrastructure of public schools to focus on what goes on inside the classroom. The highest percentage of our resources goes towards academic improvement and educational initiatives for the children; which entails teachers’ training.”²

Study Design.— Using a random number generator, we randomly assigned 607 teachers to one of the following treatment arms: (i) utilitarian treatment (121 teachers); (ii) malleability treatment (121 teachers); (iii) visual narrative treatment (122 teachers) (iv) joint visual narrative and gender-rights curriculum treatment (121 teachers) and (v) the control treatment was provided information on procedures to open a bank account in Pakistan (122 teachers). The complete transcripts of the Utilitarian and Malleability treatments are reported in Tables A1 and A2, respectively. Transcripts to structure the discussion were identical for each treatment arm, including the control arm, and are reported as Table A3 in Appendix A. Figure B1 presents a comprehensive summary of the outcomes, treatment specifics, timeline, and principal findings of this study, alongside other studies conducted in partnership with the Progressive Education Network in Pakistan. This integrated perspective aids in the comparative analysis and meta-interpretation of the current research.

Experimental Implementation Details.— Baseline, midline, and endline surveys were conducted in February, September, and March, respectively. We organized classroom-level study groups from four months before the first mathematics examination until the math finals. We cross-randomized teachers to organize students within their classes in mixed-gender or same-gender study

² Progressive Education Network (2022). PEN Mission Statement. Retrieved from: <http://www.pen.org.pk/our-approach-9464>

groups (from April to July). Figure 1 provides an illustration of the study groups during authors' random spot visits. We conducted the first round of data collection six months after the intervention (short-term effects), while the second round of data was collected 12 later. In both rounds, the data collection was supervised by school administrators, and the PEN staff. We worked with a local training department within PEN to manage the surveys with support from our research enumerators.

Utilitarian and Malleability Treatments.— This pair of two-hour interventions targeted broad-based compassion that highlighted global attentiveness to all groups, rather than being limited to gender attitudes. Specifically, the utilitarian intervention was designed to underscore the practical benefits of fostering empathy in educational settings (Mehmood, Naseer, Chen, 2024). By emphasizing the individual benefit of being empathetic, we aimed to align the teachers' self-interest with the promotion of empathy towards others and, by extension, gender-positive attitudes. Similarly, the malleability of empathy training intervention was based on the understanding that empathy is a trait that can be cultivated and developed. Such training could potentially extend to fostering more gender-progressive attitudes. These treatments hoped to leverage a common human capacity for understanding of the “outgroup” and a rational compassion, which can be impactful in shaping attitudes towards gender equality. Full transcripts of these interventions, that both hoped to shift empathy by relying on narratives and evidence, are presented in Tables A1 and A2 of Appendix A, respectively.

Visual Narrative Treatment.— Our third treatment group attended a live screening of a movie—an emotionally charged social drama lasting 2 hours and 45 minutes. The visual narrative was the 2011 film "Bol" (meaning "Speak Up"), an Urdu-language social drama co-produced by a Pakistani director Shoaib Mansoor and the Johns Hopkins University. In this film, a strong female lead,

who is on death row, explains why she found it necessary to murder her father as her “right to exist as a woman” was subverted. The movie explores how she and her sisters deal with her father’s obsession with having a son, his perpetuation and strict enforcement of regressive gender attitudes (such as limiting his daughters’ rights to employment, inheritance, education, and public spaces), and his staunch rejection of his existing intersex child. This quote from the movie sums up the theme:

“The distance between the rights and freedoms enjoyed by boys in the school next door and us girls in this house is ostensibly one of a wall of a few inches, but the real distance between us and them extends thousands of miles.”

The movie was followed by a 30 minute structured discussion on the movie with application to women’s rights and gender attitudes in society. During this discussion, the teachers empathized with the strong female character and discussed the movie’s portrayal of the gender gap in rights such as education, work, politics, going outside the home and accessing public spaces.

Joint Visual Narrative and Gender-Rights Curriculum Treatment.— Our fourth treatment combined the movie “Bol” with a teacher-training session on how to conduct a semester-long gender-rights course. This is inspired by theory and empirical evidence on the efficacy of social-emotional learning and teaching as an instrument of self-persuasion (Eskreis-Winkler et al., 2019). In this treatment, the movie was followed by a three-hour workshop on teaching a semester-long course titled “Gender Equality in Child Development for Social Skills.” The course commenced in April 2021 and spanned four months. It consisted of three-hour lectures held weekly, amounting to a total of 48 hours of classes over the four-month period. The course was designed by gender activists, the authors of this study, educators, pedagogical consultants, and multimedia developers. The teachers were instructed to discuss gender rights in a context

applicable to the child and to his immediate family, and to organize in-class exercises. The classroom exercises involved reading, drawing, and other activities encouraging students to reflect on gender rights, women's place in society, and the rights and freedoms women enjoy. Typical tasks noted by the teachers in their activity logs included:

Task 1) Draw all the work that your father does.

Task 2) Draw all the work that your mother does.

Task 3) Which of these are the same? Which are different?

Task 4) Why is that?

Further examples of lesson plans and teachers' activity logs are reported in Figure B2 of Appendix B. Within this treatment arm, a 30 minute structured discussion similar to the standalone visual narrative treatment on the Bol movie was also held.

Implementation Details.— The training was conducted live on Zoom by the same team of 8 research assistants located in Lahore, using the same set of questions, resulting in similar discussion duration across groups. The screening of the Bol movie was also conducted live on Zoom, in adherence with PEN training requirements, the teachers had their cameras on. The curriculum cost was similar as it was developed by the research team and a group of pedagogical volunteers, and the teachers did not charge any fees due our collaboration with PEN.

Control Group.—The control group receives training on generic procedures to open a bank account in Pakistan. This includes readily known facts such as going to the bank reception, requesting to open a bank account and presenting identification documents. Administrative data from Progressive Education Network reveal that every teacher in our sample already had a bank account at time of this intervention, so this control treatment is unlikely to have a direct effect on gender attitudes.

Mixed and Same Gender Study Groups.— Building on recent scholarship in pedagogy that effective cooperation may be a key pathway to improve math achievement (Gutiérrez, 2002; Weissglass, 2000; Corte, 2004), and intergroup contact reduces prejudice (Rao, 2019; Lowe, 2021), we cross-randomized teachers to form either mixed-gender or same-gender student study groups in their mathematics classes. The instructions provided to teachers were to make a group of two-students (that were randomly assigned either to be mixed or of same-gender). They were to meet once weekly for 30 minutes and discuss any past homework assignments in the assigned group of 2. This was done within the mathematics class under teachers' supervision for 3 months before the mathematics final exam. There was no further instruction for the students or teachers of either study group on the structure of the group discussion except we will do random spot visits in school to ensure enforcement.

Behavioral Games with Students.— We employ administrative data to evaluate student achievement in mathematics exams and gauge their gender attitudes through attitudinal surveys and behavioral games. By analyzing students' behavior in games that involve either same-gender or opposite-gender interactions, we seek to determine the impact of gender on students' treatment of others. To this end, we implement four incentivized economic games: competition, cooperation, coordination, and redistribution, inspired by the methodology employed in Kosse et al., 2020, in their work with primary school students. Adapting these games to fit cultural norms, we introduce the "milk bank" version where winners receive "milk carton coupons" redeemable for milk cartons at school canteens.³

³ Table A4 details Research Ethics, whereas Tables A5 and A6 discuss curriculum treatment and teacher workshops, respectively.

III. Data and Empirical Strategy

A. The Data

Sample.— The sample consists of all 607 teachers and their 13,911 students across all 52 schools chartered by PEN in the State of Punjab. As is common in most public schools in Pakistan, all teachers are female and teach every class from Kindergarten to Grade 6. The students, however, are of mixed-gender in public schools of Pakistan until they ‘graduate’ from Grade 6. Nevertheless, this is not a universal practice. Our sample consists of 7107 boys and 6804 girls ranging from age 5 to 12. The PEN network organizes several training workshops for teachers each year, and our experiment took place within the PEN teacher training drives. As a result of our experimental intervention embedded within the PEN’s regular training programs, we essentially have zero attrition. All 607 PEN school teachers in the State of Punjab participated in the experiment. The baseline survey was carried out in the second week of January, the midline survey in August, and the endline in March 2022. We utilize detailed administrative data on teachers and students, enabling us to match teachers to their students at the classroom level.

Outcome Variables on Teachers’ Gender Attitudes.— Our first set of outcome variables concerns teachers’ attitudes towards gender rights as assessed about six months and a year after our intervention. To summarize teachers’ gender attitudes, we use the gender rights index that averages across all components of gender rights survey questions listed in Appendix D. Women’s Rights Overall is an index combining all the statements concerning women’s economic, social, legal, and political rights. Women’s Economic Rights is an index combining women rights regarding education and work. Women’s Political Rights is based

on the right to hold political office, while Women's Social Rights and Legal Rights concern the rights to choose whether or not to conform to social and legal discrimination, respectively. For more details, see our survey instrument in Appendix D2, while information on index construction is reported in Appendix D3.

Outcome Variables on Teachers' Decisions and IATs.— Our second set of outcomes involves revealed preference measures of gender attitudes in the form of teachers' willingness to sign and send petitions asking parliament to repeal discriminatory laws, and the Implicit Association Test (IAT). The gender IAT measures implicit associations regarding women. The IAT involves categorizing words by placing them on the left or right of a computer screen and measures the strength of association between two concepts based on response times. We use a standard Gender-Career IAT test to see if respondents associate women with "Family" and men with "Career"). The IAT hence measures female-sounding names and gender stereotypes.⁴ The gender that we administered IAT was the standard career-family word association task based on 7 questions. Too short or long latencies were automatically dropped according to the algorithm determined in Greenwald et al., (2009).⁵

Outcome Variables on Students.— We compute students' gender attitudes by fielding a 5 statement survey, 6 and 12 months following the treatment. The survey statements are chosen so that it may be understood by primary school students of different ages that populate our sample. The survey statements are

⁴ The use of IATs also reduces concerns about experimental demand. First, it is hard to respond to IATs in a socially-desirable way, as this would require strategically speeding up or slowing down in certain blocks of associations (Alesina et al., 2018). Second, consistent with the psychological literature, our IAT algorithm discards observations that are too slow or too fast (Greenwald et al., 2003). Last, it is highly unlikely that teachers within our sample—primary and middle school teachers in Punjab—would know about IATs.

⁵ We administer the IAT in Urdu online over Otree and report the text that the teachers saw on their screens in Appendix D5, while exact template for petition text is provided in Appendix D4.

reported in Appendix D6. We are also able to obtain results on standardized mathematics examinations that the students gave, also 6 and 12 months following the treatment. We use these outcomes to assess both students’ gender attitudes and academic achievement. We also play a total of four games with the students that include cooperation, coordination, redistribution and competition. Following Kosse et al., 2020, we adapt the standard games into a “piggy bank” version which we call the “milk bank” game due to cultural reasons. In particular, we offer winners “milk carton coupons” that could be redeemed at the school canteens for milk cartons. More details on the games, including the transcripts the students saw can be found in Appendix D7.

Main Explanatory Variables.— Our key explanatory variables are dummies for the four treatments. U_i and M_i denote dummies that switch on if the teachers were assigned to the Utilitarian or Malleability treatments, respectively. $Bol\ Movie_i$ (BM) and $Bol\ Movie\ \&\ Curriculum_i$ (BMC) are dummies that switch on if the teachers were assigned to the Visual Narrative or Joint Visual Narrative and Gender-rights Curriculum treatments, respectively. The control group receives training on generic procedures to open a bank account in Pakistan. This includes readily known facts such as going to the bank reception, requesting to open a bank account and presenting identification documents.⁶

B. Attrition and Balance

Close cooperation and support from the leadership of the PEN organization on administrative data and the fact that the experiment was embedded within PEN’s regular trainings meant that attrition was zero for teachers (except the attrition artificially triggered by the IAT algorithm for IAT

⁶ Administrative data from Progressive Education Network reveal that every teacher in our sample already had a bank account at time of this intervention, so this control treatment is unlikely to have direct effect on outcomes.

scores), and student attrition amounted to only 21 students in our midline and endline surveys (held about 6 months apart).⁷ Nevertheless, a lack of balance might still complicate causal interpretation of our results. We therefore examine whether our randomization was successful in creating balance among teachers and students. Table B1.2 in Appendix B shows individual characteristics, with Panel A reporting the treatment balance over teacher characteristics and Panel B regarding student characteristics. Differences across treatment groups are small in magnitude, and almost all estimated p-values are larger than 0.10, suggesting that the randomization was effective at creating balance between the groups. For instance, teachers' education, experience, class size, number of hours of teaching, marital status, and pre-treatment gender attitudes are balanced across treatment and control groups. Likewise, from Panel B of Table B1.2, we observe that the pre-treatment gender attitudes and mathematics test scores of students are balanced. Table B1.1 presents the descriptive statistics of the main outcomes used in the study. Tables B1.2 and B1.3 provide evidence of balance for the baseline specifications, while Table B1.4 also demonstrates balance for our cross-randomized school study groups in both same-gender and mixed-gender groups. To ease comparisons with baseline estimates, we also report results in standardized units. In all these instances, the treatments appear to be balanced across teacher and student characteristics.

C. Estimation Strategy

The impact of our four treatments can be evaluated by comparing outcomes across groups in a simple regression framework. For each outcome, the estimation equation is:

⁷ This remarkable take-up was only possible due to gracious support and cooperation of the Director of Training and Research, Miss Sumera Morris and her staff at PEN. They provided invaluable suggestions and support throughout this intervention.

$$Y_i = \alpha + \beta U_i + \gamma M_i + \delta BM_i + \omega BMC_i + \theta_s + X_i \mu + \epsilon_i$$

(1)

where Y_i is the outcome for a teacher i , U_i is a dummy variable equal to one if the teacher is assigned to the utilitarian empathy treatment; M_i is a dummy variable equal to one if the teacher is assigned to the malleability empathy treatment; BM_i is a dummy variable equal to one if the teacher is assigned to the visual narrative (movie “Bol” promoting more equitable gender rights) treatment and BMC_i if the teacher is assigned to the joint visual narrative and gender-rights curriculum treatment. θ_s represents school fixed effects. X_i is a vector of individual-level teacher and student controls. In equation (1), β measures the effect of the utilitarian treatment; γ the effect of the malleability treatment; δ the effect of the visual narrative treatment; and ω measures the effect of the joint visual narrative and gender-rights curriculum treatment. Following Chetty et al., 2014, in student-level regressions, we also always control for the student’s prior test scores. We cluster standard errors at the teacher level for students since that is our level of randomization. Since randomization is at the individual level, so Newey-West robust p-values are included as baseline.

In addressing the inherent challenges of multiple hypothesis testing (MHT) and the associated risk of inflated Type I errors, our empirical analysis employs two statistical methods to control the False Discovery Rate (FDR) and the Familywise Error Rate (FWER). We adopt Anderson's (2008) two-stage refinement of the Benjamini-Hochberg FDR procedure, which sequentially adjusts significance levels to reduce false discoveries. For FWER control, we apply the Romano-Wolf step-down procedure via the `rwolf2` package, leveraging its resampling mechanism to accommodate test correlations. All main tables

include Newey-West standard p-values as well as FDR q-values and FWER-adjusted p-values.

IV. Impact on Teachers

Effects on Teachers' Attitudes.— We measure teachers' attitudes 12 months following the treatment and observe quantitatively and qualitatively significant impact of visual narrative treatments, both on its own and when combined with the semester-long gender-rights curriculum. Table 1 reports the results on the impact of all our treatments on gender attitudes, while Figure 1 visualizes all of these impacts, including the stand-alone visual narrative treatment and the impact of the joint visual narrative and curriculum treatment. Figure 1 presents the estimated coefficients for the medium-term impact, measured at 12 months post-treatment. For the short-term effects, observed at a six-month interval after the treatment, refer to Table C1 in Appendix C. From Column 1 of Table 1, we observe that the visual narrative alone increased support for more equitable gender rights by about 0.15 standard deviations. When the visual narrative is combined with the curriculum, the impact almost doubles, with teachers' support for women's rights increasing by roughly 0.20 standard deviations. The effects are enduring, evident in both the short-term (as shown in Table C1) and medium-term (as indicated in Table 1), with the robustness of the results confirmed by Newey-West p-values, Sharpened q-values adjusted for multiple hypotheses, and Romano-Wolf correction for Family-Wise Error Rate (FWER) p-values. Table 1 also contains results on the impact of the Utilitarian and Malleability treatments: these do not appear to have much impact on gender attitudes either in the short or medium term. These overall gender attitudes index is also disaggregated into economic, political, social, and legal rights indices. The coefficient estimates suggest that the visual narrative and joint treatments likely shifted attitudes concerning women's economic, political and legal rights. As a

point of comparison, the impact of the joint treatment of visual narrative and 4-month gender studies curriculum is roughly half the effect size as that found for random roommate assignment (Corno et al., 2022) and a little over half the effect as that found for a 12-week training program (Devine et al. 2012).

Effect on Teachers' IATs.— Next, we assess the impact of the standalone visual narrative and the joint visual narrative and curriculum treatments on the Implicit Association Test (IAT) scores. From Figure 2, we observe that the visual narrative treatment alone reduces implicit gender bias by about 0.20 standard deviations, and when it is combined with the gender-rights curriculum, implicit gender bias reduces by at least 0.32 standard deviations. These effects hold for both the short (Panel A) and mid-term (Panel B) and when we consider the IAT in original units or raw scale (Figure B3 in Appendix B). We interpret the greater effect of the joint treatment as a sign that teaching the gender-rights curriculum reinforced the treatment through self-persuasion, consistent with recent evidence that even short amounts of time spent discussing principles can result in self-persuasion (Eigen and Listokin, 2012; Schwarzmann, Tripodi, and van der Weele, 2022).

Effects on Teachers' Decisions (petitions).— Finally, we provide causal evidence that self-reported attitudes are reflective of behavioral change in high-stakes settings. We offer all the teachers the opportunity to sign a petition that is then sent to the Pakistani parliament seeking the abolishment of discriminatory laws allowing polygamy, as well as a petition seeking to make a man's demand for a dowry a criminal offense. The results here are even more striking than what we found for self-reported gender attitudes. From Table 2, we observe that the visual narrative of the movie “Bol” alone increases the likelihood of petitions sent by teachers by about 10 percentage points. This is economically significant and equivalent to roughly a doubling of petitions over the sample mean. The coefficient estimates for the joint treatment are even larger, with

petition-signing increasing by about 20 percentage points. These results strongly suggest that our results from self-reported surveys are likely to have real-world implications.

Effect of Previously Seeing Bol Movie.— About half of the teachers stated that they had watched the Bol movie after it was released a decade earlier in 2011, therefore, we investigate if those teachers who had previously watched Bol are more likely to be impacted by the visual narrative or the joint treatment, notwithstanding the endogeneity of the teachers' prior exposure to the movie. Figure B4 shows little evidence for the heterogeneous effect of treatment on those who had previously watched the movie Bol. These results hold for gender attitude surveys, IAT and the two gender rights petitions. We first assess whether exposure to the film "Bol" influences the attitudes of control group teachers, finding no association with our measured outcomes (Figure B4, Appendix B). Subsequently, we analyze treatment effects among teachers who had previously seen the movie, confirming that having watched the "Bol" movie is not correlated with gender attitudes in the control group (Figure B5). Lastly, we demonstrate consistent effects of the Visual Narrative and Joint Treatment on teachers who had previously seen the "Bol" movie; these teachers also exhibit a similar change in gender attitudes (Figure B6 and Table B10). Taken together, the results suggest that watching the "Bol" movie beforehand is unlikely to overturn our findings. This may be due to many factors. For instance, if the teachers watched the movie a decade earlier when it premiered, and the effects of the movie on teachers may have dissipated after 10 years. It could also be the case that the structured discussion on the gender rights themes of the movie among peers reinforced the message of the movie beyond just watching the movie.

V. Impact on Students

Effects on Students' Attitudes.— To investigate whether gender-rights attitudes were transmitted from teachers to students, we surveyed all students that the 607 teachers taught, roughly 7 months and one year later. Columns 1 and 2 of Table 3 report these results. We observe that the visual narrative alone made students more supportive of gender rights by about 0.10 standard deviations. The results hold for both the short and medium term, suggesting that gender-rights transmission is likely persistent. These results are also particularly striking since the standalone visual narrative treatment *squarely* focused on the teachers. Nevertheless, consistent with results on teachers, the coefficient estimates from Table 3 also indicate more than a doubling of the effects under the joint visual narrative and curriculum treatment: students whose teachers both experienced the visual narrative and taught the gender-rights curriculum are at least 0.25 standard deviations more supportive of more equitable gender rights in both the short and the medium term.

Effects on Students' Test Scores.— One of our treatments also impacts student achievement, as shown by the students' mathematics test scores. We observe that math test scores are positively impacted when the teachers were assigned to the joint visual narrative and curriculum treatment. These results also allay potential concerns of experimental demand since the treatment focused on gender attitudes, not math achievement. Columns 3 and 4 of Table 3 report these results: students' math scores are about 0.1 standard deviations higher under the joint treatment. To put this into perspective, the observed 0.1 standard deviation rise in mathematics test scores implies a 3.5-point enhancement on the 0 to 100 grading scale used for the math assessment. This improvement reflects a 5% increase above the average score of the sample of students in our study or student outcomes for grade B+ students increase to A- due to our treatment. While the

standalone Bol movie treatment of teachers does not affect student test scores, gender-focused interventions, such as film screenings joint with a curriculum, can shift gender attitudes and even student test scores, indicating that more intense interventions are needed to alter student test scores.

It is also worth noting, that the smaller and statistically insignificant point estimates of the standalone movie treatment on academic test scores do not support the hypothesis that exposure to the visual narrative, specifically the Bol movie, independently contributes to improvements in academic test scores. This finding implies that interventions of a light-touch nature, which are limited to teacher engagement, are not potent enough to elicit changes in the academic performance of students. In the next section, we investigate a mechanism behind this increase, deploying an experimental causal mediation design, to investigate how students learned more effectively in the joint treatment arm.

Effect on Stress of Students.—In a separate analysis of the same intervention, Mehmood, Naseer, and Chen (2023) found that while the visual narrative treatment arms succeeded in shifting teacher’s gender attitudes, they adversely impacted mental health and had adverse consequences in terms of domestic violence, though the negative effects on mental health disappeared if a large enough share of teachers in the school were also exposed to the visual narrative treatment, a phenomenon we call “moral bandwagoning”. It is important to note that the increase in teacher stress has not led to a decline in student academic performance. Access to national mathematics exam results allows us to determine that the implementation of our visual narrative, alone or in conjunction with other treatments, has not adversely impacted academic achievement in the short run (6 months after treatment) or in a more extended period (12 months after the intervention). We further observe that while teacher stress levels rose in response to the visual narrative and the combined visual narrative with curriculum treatment, this pattern was not mirrored among students. Student stress levels are

statistically similar across the treated groups when compared to the control group. These results are reported in Table B3 in Appendix B.⁸

Our results remain robust across numerous sensitivity tests, as detailed in a robustness section presented in Appendix A2. Here, we address potential concerns such as experimental demand and spillover effects, concluding that our results are unlikely to be influenced by these factors. This conclusion is also supported by the use of the Implicit Association Test and the Marlowe-Crowne social desirability scale, coupled with observed the sustained effects of our treatments over time.

VI. Mediation Analysis

The design of our experiment lends itself to causal mediation analysis to study inter-gender contact driving our results. Typically, a sequence of behavioral data is observed to infer how earlier actions mediate final outcomes. Here, we use an experimental set-up of randomization across teachers and re-randomization of teachers to form mixed-gender or same-gender student study groups to causally isolate the effect of working with the opposite gender up to one year after the training intervention. The results of our causal mediation analysis help uncover experimentally why students' test scores increased for the joint visual narrative and curriculum treated group.

Building on the emerging causal evidence that intergroup contact may reduce prejudice (Rao, 2019; Lowe, 2021), we hypothesized that more equitable gender-rights attitudes would be formed through interactions with the opposite gender, and that these interactions would lead to teacher transmission of equitable

⁸ Further heterogeneity analysis reveals that the impact of visual narrative and joint treatment on gender attitudes remains relatively uniform across various grades of students. Nonetheless, the point estimates reveal that a combined intervention might exert a greater influence. These findings are detailed in Table B4, located in Appendix B. A similar heterogeneity analysis by gender is also performed, reported in Table B5, with the results being similar between boys (Panel A) and girls (Panel B).

gender-rights attitudes to students, and improved theory of mind via increased cooperation and coordination with the opposite gender. Two months before the math exam, we randomly assigned teachers to organize students into mixed-gender or same-gender study groups that lasted until the final examination held about a year after the treatment. Random spot checks by the authors, field assistants, and PEN administrators confirmed that our study group randomization was followed by the teachers.

Effects on Students' Math Scores by Study Group.— The increase in test scores is driven by students whose teachers were assigned to the joint visual narrative and curriculum treatment and those rerandomized to form opposite-gender study groups. The math achievement of students assigned to same-gender study groups is unaffected. The estimates from Column 1 of Table 4 indicate that students randomly assigned to teachers treated with the joint visual narrative and gender-rights curriculum and opposite-gender study groups increased their test scores by roughly 0.18 standard deviations.⁹ The impact persists over both the short and the medium term. We observe similar results from an interaction specification (Table 4, Column 5): there is a qualitatively and quantitatively significant impact on math scores when the teachers treated with the joint visual narrative and gender-rights curriculum treatment formed the mixed-gender study groups. Figure B7 gives raw mathematics scores out of 100, revealing the same impact. Overall, the results presented in Table 4 and Figure B7 indicate that de-segregation within classes may enhance learning. Importantly, de-segregation has no independent effect in the control group.

Effects on Students' Math Scores by Gender.— Intriguingly, there is a positive impact on math achievement for *both* boys and girls (Figure B8). The male student scores increase by about 0.15 standard deviations, while the female

⁹ The short-run results, observed six months post-treatment, are consistent with medium-term findings and can be reviewed in Table C3 of Appendix C.

scores increase by about 0.1 standard deviation, both in the short- and medium-term. The mathematics achievement gender gap decreases by at least 20%. Figure B8 shows that both boys' and girls' overall math scores increase from about 65 to 70 out of 100, following our joint narrative and curriculum treatment, i.e. student outcomes improve from grade B+ to A- for the treated group. The greater improvement in test scores for boys, as compared to girls, may suggest that the joint narrative and curriculum treatment is particularly effective in addressing the learning needs of male students, it is likely that since boys had lower baseline achievement levels in mathematics at an early age compared to girls, the treatment might have had more room for impact, leading to a larger gain for boys as they "catch up" to their female peers. This is consistent with a recent meta-analysis performed by Kersey et al. (2018), which synthesized data from six studies to explore gender differences in early mathematical cognition among children. The findings indicated that there were no significant differences in mean performance and variability in mathematical abilities between boys and girls at these early stages. If anything, girls' scores in the early years were slightly higher than those of boys.¹⁰ This mirrors our case (where children aged from 6 to 12-year-old are studied), indicating the treatment could facilitate greater gains for boys as they "catch up" to girls.

Effects on Students' Behavioral Games by Study Group.— Table 5 separates the results of the behavioral games according to whether students played against the opposite gender or against the same gender (Table B2). Students playing against the opposite gender showed greater cooperation and coordination in the behavioral games. Both the visual narrative treatment and the joint treatment positively impact cooperation and coordination in strategic dilemmas

¹⁰ Similar results are found in the studies by Felson and Trudeau (1991) on gender differences in mathematics performance between girls and boys, in Pomerantz et al. (2002), Duckworth and Seligman (2006) and Hyde et al. (2008).

commonly used to measure social preferences of children. The coefficient estimates imply that the visual narrative alone increased cooperation and coordination by about 0.15 standard deviations, while the joint treatment increased cooperation and coordination by about 0.3 standard deviations (Table 5). Redistribution and competitiveness are unaffected. The behavior of students playing the strategic dilemma against the same gender (Table B2) also does not seem to be impacted by any of our treatments, while Figure B9 provides corresponding raw averages of the four treatments on games fielded with the opposite gender and Figure B10 provides these effects by mixed gender and same gender study groups. Taken together, these findings suggest that the gender-rights curriculum combined with mixed-gender study groups may have enhanced students' theory of mind towards the opposite gender – the ability to take another person's perspective. More gender-equal attitudes may have arisen through interactions with the opposite gender, promoting both theory of mind and transmission of progressive gender attitudes. The two reduced form effects on cooperation and on student achievement corroborates a body of scholarship that advocates cooperative learning enhances mathematics learning (Gutiérrez, 2002; Weissglass, 2000; Corte, 2004). These findings suggest a potential mechanism explaining why inter-gender contact may improve mathematics achievement. Effective cooperation across genders may yield benefits in student achievement.

Comparison with studies of single vs. mixed gender schools.— We observe positive effects of cross-gender interaction in Pakistani primary schools, a finding that stands in contrast to Briole (2021), who found in France that while a higher proportion of female peers boosts girls' academic performance and future educational success, it concurrently steers boys towards vocational paths and diminishes their likelihood of graduating. Likewise, Jackson (2021) also reported benefits from transitioning coeducational schools to single-sex in Trinidad and Tobago, noting improved exam results for boys, more rigorous coursework for

both sexes and declines in arrests and teen pregnancies. Our study's distinct context, focusing on primary education for children aged 5 to 12, and explicit focus on cultivating progressive gender norms in a child's formative years may explain these different results. During these critical early years, implementing forward-thinking gender interventions can nurture positive attitudes toward the opposite gender and effectively challenge the harmful behaviors and stereotypes that may be observed in earlier work. Furthermore, the students in the early teenage years analyzed in the studies mentioned previously may have already been affected by established social norms that discourage cross-gender interaction. Consequently, concentrating on a younger age group that may be more receptive to increasing inter-gender cooperation may also account for the differing outcomes observed in our study.

VII. Conclusions

The past century has witnessed tremendous growth in recognition of rights and freedoms across group boundaries. This paper explores how teachers may transmit gender attitudes to their students, through a field experiment in Pakistan. We implement a randomized control trial testing different methods of shifting teachers' and students' views regarding equitable gender rights. We field teacher-training treatments based on the utility of empathy, malleability of empathy, a visual narrative arguing for female rights, and a joint visual narrative and gender-rights curriculum, as well as a control training.

We find that training teachers using a visual narrative shifted the teacher's attitudes towards more equitable gender rights. Teacher attitudes impacted both student attitudes and students' behavior towards the opposite gender. The effect sizes are substantial. Teachers' attitudes measured in gender IATs shifted by 0.2 standard deviation. The teachers also became 10 percentage points more likely to

sign a petition sent to the Pakistani parliament to criminalize mens' demands for a dowry and abolish polygamy laws. Students' attitudes shifted towards gender equality by at least 0.1 standard deviation. Reinforcing the visual narrative shown to teachers with the gender-rights curriculum improved student achievement and magnified the shifts in students' attitudes. Teachers assigned to the joint visual narrative and gender-rights curriculum treatment shifted their own attitudes and behavior by an additional 0.35 standard deviations in IAT scores; they were 18 percentage points more likely to sign gender rights petitions sent to the parliament. We interpret this heightening of treatment effects by teachers exercising self-persuasion via teaching. Students' attitudes on more equitable gender rights, students' behavior towards the opposite gender in games reflecting theory-of-mind, and students' math test scores were all also impacted under the joint visual narrative and curriculum treatment.

Our causal mediation design experimentally identifies a mechanism explaining the increase in student achievement. Cross-randomizing teachers to assign students into mixed-gender or same-gender study groups shifts students' attitudes. Transmission of attitudes occurred only when classrooms were assigned to mixed-gender study groups, as reflected in behavioral games and math scores. Inter-gender cooperation and coordination increased in games involving social dilemmas, suggesting improved theory of mind regarding others' actions.

The gender gap in math test scores essentially disappeared in classrooms assigned to form mixed-gender study groups, likely due to the treatment increasing cooperation and coordination with the opposite gender. Overall, our results provide experimental evidence that gender attitudes are transmissible from teachers to students and underscore the potential benefits of contact with the opposite gender in learning environments.

While gender separation of school children in Pakistan is a potential practice post grade 6 education, it is by no means a standard across all educational

institutions in Pakistan, with many private schools favoring a mixed-gender education. Although, we have documented the immediate and intermediate effects of gender separation, it is important to consider the possibility that these effects may not be enduring or could reverse in the context of higher education, particularly when the practice of mixed-gender schooling is not sustained. A more complete understanding long-term impact of gender separation in education merits additional scrutiny.

References

1. Andrabi, T., Bau, N., Das, J. & Khwaja, A.I. (2024) Crowding in Private Quality: The Equilibrium Effects of Public Spending in Education. Working Paper. Quarterly Journal of Economics.
2. Alan, S., Baysan, C., Gumren, M. and Kubilay, E., 2021. Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, 136(4), pp.2147-2194.
3. Alesina, A., Giuliano, P. and Nunn, N., 2013. On the origins of gender roles: Women and the plough. *The quarterly journal of economics*, 128(2), pp.469-530.
4. Alesina, A., Carlana, M., La Ferrara, E. and Pinotti, P., 2018. *Revealing stereotypes: Evidence from immigrants in schools* (No. w25333). National Bureau of Economic Research.
5. Banerjee, A., Ferrara, E.L. and Orozco, V., 2019, May. Entertainment, education, and attitudes toward domestic violence. In *AEA Papers and Proceedings* (Vol. 109, pp. 133-37).
6. Bénabou, R., Falk, A. and Tirole, J., 2018. *Narratives, imperatives, and moral reasoning* (No. w24798). National Bureau of Economic Research.
7. Bisin, A. and Verdier, T., 2011. The economics of cultural transmission and socialization. In *Handbook of social economics* (Vol. 1, pp. 339-416). North-Holland.
8. Briole, S., 2021. Are girls always good for boys? Short and long term effects of school peers' gender. *Economics of Education Review*, 84, p.102150.
9. Chetty, R., Friedman, J.N. and Rockoff, J.E., 2014. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), pp.2593-2632.

10. Corno, L., La Ferrara, E. and Burns, J., 2022. Interaction, stereotypes, and performance: Evidence from South Africa. *American Economic Review*, 112(12), pp.3848-3875.
11. Corte, ED, 2004. Mainstreams and perspectives in research on learning (mathematics) from instruction. *Applied Psychology* , 53 (2), pp.279-310.
12. Crowne, D.P. and Marlowe, D., 1960. A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, 24(4), p.349.
13. Devine, P.G., Forscher, P.S., Austin, A.J. and Cox, W.T., 2012. Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of experimental social psychology*, 48(6), pp.1267-1278.
14. Dhar, D., Jain, T. and Jayachandran, S., 2022. Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American economic review*, 112(3), pp.899-927.
15. Doepke, M. and Tertilt, M., 2009. Women's Liberation: What's in it for Men?. *The Quarterly Journal of Economics*, 124(4), pp.1541-1591.
16. Doepke, M. and Zilibotti, F., 2017. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5), pp.1331-1371.
17. Duckworth, A. L., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98(1), 198-208.
18. Duflo, E., 2012. Women empowerment and economic development. *Journal of Economic literature*, 50(4), pp.1051-79.
19. Eigen, Z.J. and Listokin, Y., 2012. Do lawyers really believe their own hype, and should they? A natural experiment. *The Journal of Legal Studies*, 41(2), pp.239-267.

20. Eskreis-Winkler, L., Milkman, K.L., Gromet, D.M. and Duckworth, A.L., 2019. A large-scale field experiment shows giving advice improves academic outcomes for the advisor. *Proceedings of the National Academy of Sciences*, 116(30), pp.14808-14810.
21. Felson, R. B., & Trudeau, L. (1991). Gender differences in mathematics performance. *Social Psychology Quarterly*, 54(2), 113-126.
22. Fernández, R., 2014. Women's rights and development. *Journal of Economic Growth*, 19(1), pp.37-80.
23. Fernández, R. and Wong, J., 2014. Unilateral divorce, the decreasing gender gap, and married women's labor force participation. *American Economic Review*, 104(5), pp.342-47.
24. Field, E., Pande, R., Rigol, N., Schaner, S. and Troyer Moore, C., 2021. On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms. *American Economic Review*, 111(7), pp.2342-75.
25. Giuliano, P., 2020. Gender and culture. *Oxford Review of Economic Policy*, 36(4), pp.944-961.
26. Giuliano, P. and Nunn, N., 2021. Understanding cultural persistence and change. *The Review of Economic Studies*, 88(4), pp.1541-1581.
27. Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L. and Banaji, M.R., 2009. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology*, 97(1), p.17.
28. Greenwald, A.G., Nosek, B.A. and Banaji, M.R., 2003. Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of personality and social psychology*, 85(2), p.197.
29. Gutiérrez, K.D., 2002. Studying cultural practices in urban learning communities. *Human development*, 45(4), pp.312-321.

30. Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321(5888), 494-495.
31. Isensee, L. (2016) 'US Charter School Model Goes Global in Pakistan', Houston Public Media, 29 November. Available at: <https://www.houstonpublicmedia.org/articles/news/education-news/2016/1/29/178120/us-charter-school-model-goes-global-in-pakistan/> (Accessed: 5 June 2024).
32. Jackson, C.K., 2021. Can introducing single-sex education into low-performing schools improve academics, arrests, and teen motherhood?. *Journal of Human Resources*, 56(1), pp.1-39.
33. John Hopkins, 2012. Film Co-Produced by JHU-CCP Wins at Pakistani Awards Show By Johns Hopkins Center for Communication Programs. Available at <https://ccp.jhu.edu/2012/08/03/film-co-produced-by-jhu%E2%88%99ccp-sweeps-prestigious-pakistani-awards-show/>
34. Kersey, A.J., Braham, E.J., Csumitta, K.D., Libertus, M.E. and Cantlon, J.F., 2018. No intrinsic gender differences in children's earliest numerical abilities. *npj Science of Learning*, 3(1), p.12.
35. Kling, J.R., Liebman, J.B., Katz, L.F. and Sanbonmatsu, L., 2004. Moving to opportunity and tranquility: Neighborhood effects on adult economic self-sufficiency and health from a randomized housing voucher experiment. *Available at SSRN 588942*.
36. Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H. and Falk, A., 2020. The formation of prosociality: causal evidence on the role of social environment. *Journal of Political Economy*, 128(2), pp.434-467.
37. List, J.A., Shaikh, A.M. and Xu, Y., 2019. Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4), pp.773-793.

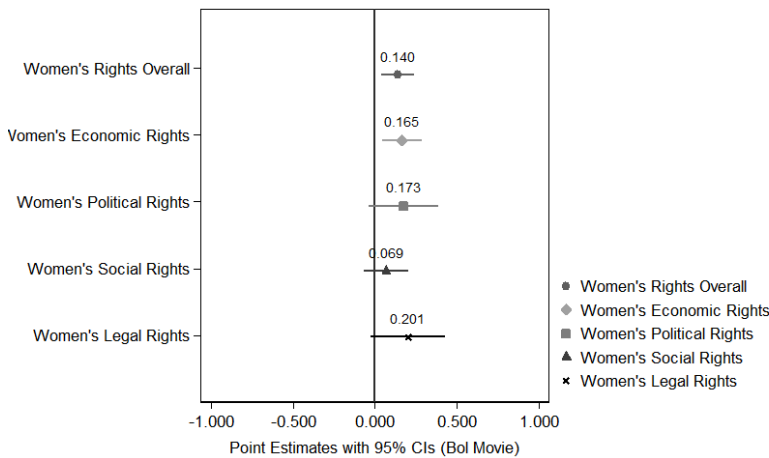
38. Lowe, M., 2021. Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6), pp.1807-44.
39. Maurin, E. and McNally, S., 2008. Vive la revolution! Long-term educational returns of 1968 to the angry students. *Journal of Labor Economics*, 26(1), pp.1-33.
40. Mehmood, S., Naseer, S., and Chen, D. 2023. Why Rights Revolutions are Rare? *Mimeo*.
41. Mehmood, S., Naseer, S., and Chen, D. 2024. Altruism in Governance: Insights from Randomized Training for Pakistan's Junior Ministers *Journal of Development Economics*.
42. National Center for Education Statistics (2019) "Digest of Education Statistics," https://nces.ed.gov/programs/digest/d19/tables/dt19_105.50.asp 2019. Accessed: 2024-05-1.
43. Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology*, 94(2), 396-409.
44. Progressive Education Network, 2022. *PEN Mission Statement*. Available at: <http://www.pen.org.pk/our-approach-9464> (Accessed: 2nd February, 2022)
45. Riley, E., 2019. Role models in movies: the impact of Queen of Katwe on students' educational attainment. *The Review of Economics and Statistics*, pp.1-48.
46. Rao, G., 2019. Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review*, 109(3), pp.774-809.
47. Seror, A., 2022. Child Development in Parent-Child Interactions, Forthcoming, *Journal of Political Economy*.

48. Schwardmann, P., Tripodi, E. and Van der Weele, J.J., 2022. Self-Persuasion: Evidence from Field Experiments at International Debating Competitions. *American Economic Review*.
49. Weissglass, J., 2000. No compromise on equity in mathematics education: Developing an infrastructure. In *In W. Secada (Ed.), Changing the faces of mathematics: Perspectives on multiculturalism and gender equity*. Reston, VA: National Council of Teachers of Mathematics.
50. Weisz, E. and Zaki, J., 2017. Empathy building interventions: A review of existing work and suggestions for future directions. *The Oxford handbook of compassion science*, pp.205-217.

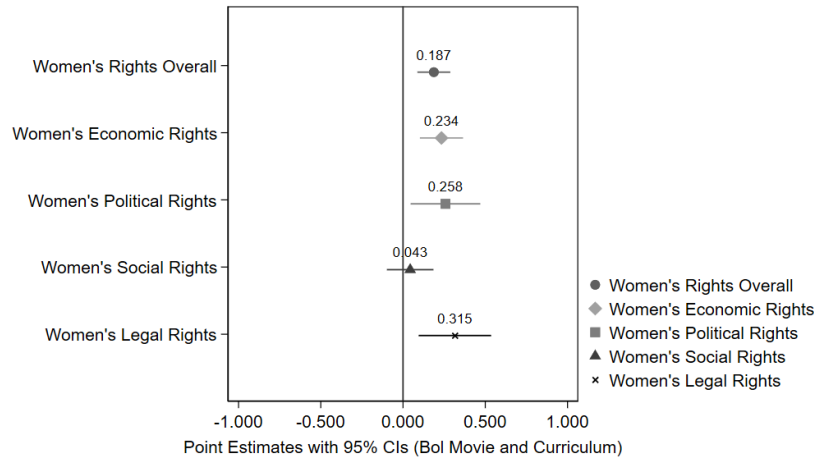
FIGURES AND TABLES

Figure 1: Impact of Movie and Joint Movie-Curriculum Treatment on Teachers' Gender Attitudes

Panel A: Bol Movie Effect



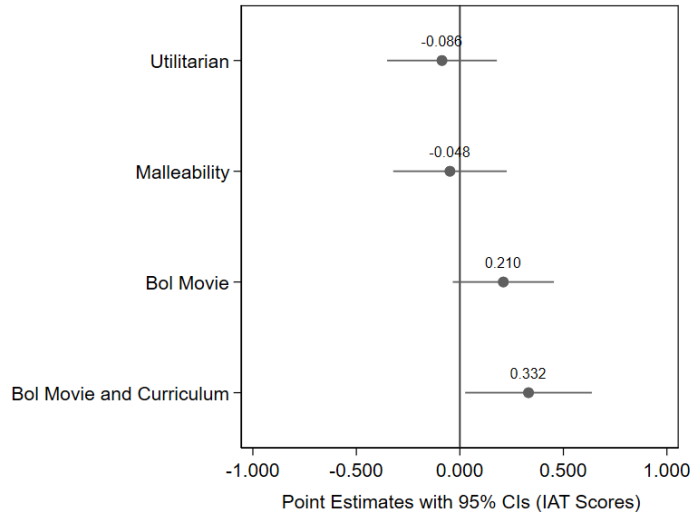
Panel B: Joint Treatment



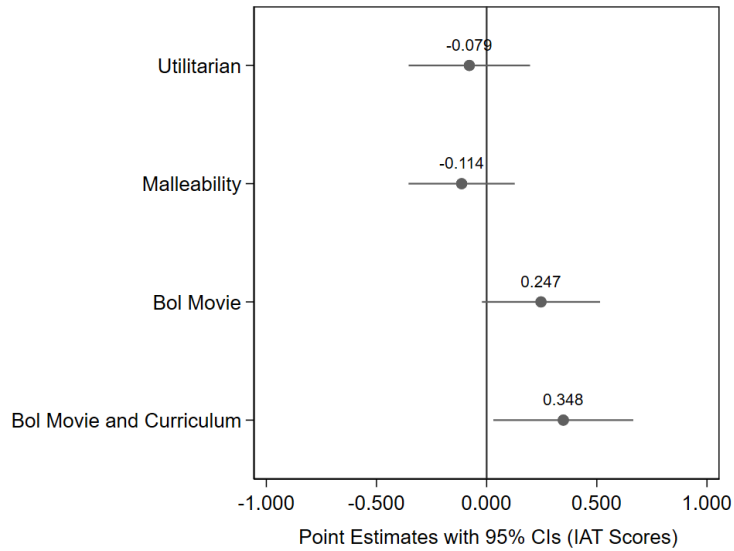
Note: The dependent variables in the figure are the gender rights indices that were constructed from survey questions listed in Appendix D. "Women's Rights Overall" is an average of all the statements concerning women's economic, social, legal, and political rights. Women's Economic Rights is an index combining women's rights to education and work outside home, based on answer to the statements "Women should be allowed to work outside the home". "Women and men should have equal rights to jobs". "I have no problem with my sister or female cousin working outside the home". "Daughters should have the same right to inherit property as sons". "Women and men should have equal rights to get an education". "Wives should not be less educated than their husbands". "Boys should not have more opportunities and resources for education than girls.". Women's Political Rights is based on statements "It would be a good idea to elect a woman as the village Sarpanch (local politician)." . "Women and men have equal rights to be President or Prime Minister.". Women's Social Rights is based on statements "Domestic violence by husbands cannot be justified" "Parents should seek their daughter's consent before fixing her marriage". "A woman should not necessarily get married before her 25th Birthday". "Women who give birth to a son need not be honored in the family". "A woman with five daughters should not be under social pressure to bear a son.". Finally, the Women's Legal Rights index is based on statements "Laws should be passed to ban dowry.". Under Article 35 of the Constitution of Pakistan & Judgment of Federal Shariat Court, the consent of 'Wali' is not required and a sui juris Muslim female can enter into a valid Nikah / Marriage under her own free will without the consent of Wali. To what extent do you approve of this legal right of women to enter marriage under their own free will." Equation (1) is estimated with all controls, but the coefficient estimate corresponding to the Movie treatment is displayed in Panel A, and the Joint Movie-Curriculum treatment is displayed in Panel B of the figure. Effects are measured 12 months post-treatment. The treatments are compared relative to the placebo treated control group.

Figure 2: Impact on Standardized Implicit Association Test scores

Panel A: Short-term Effects



Panel B: Medium-term Effects



Note: The figure reports estimates from equation (1) with all main explanatory variables and controls with gender IAT—standardized to mean zero and standard deviation one—as the dependent variable. Controls include all available individual characteristics and school fixed effects. 95% confidence bands are also reported. Short-term effects in Panel A are recorded 6 months after the treatment and medium-term effects in Panel B are recorded twelve months after the treatment. The treatments are compared relative to the placebo treated control group.

Table 1: Impact on Teachers' Attitude

	<i>Gender Rights Overall</i>	<i>Economic Rights</i>	<i>Political Rights</i>	<i>Social Rights</i>	<i>Legal Rights</i>
	(1)	(2)	(3)	(4)	(5)
<u><i>Visual narrative (movie)</i></u>	0.140	0.165	0.173	0.0687	0.201
p-value	(0.0062)***	(0.0082)***	(0.1102)	(0.3226)	(0.0838)*
Sharpened q-value	[0.022]**	[0.026]**	[0.166]	[0.303]	[0.141]
Romano-Wolf corrected p-value	{0.030}**	{0.026}**	{0.5864}	{0.9880}	{0.061}*
<u><i>Joint Movie-Curriculum</i></u>	0.187	0.234	0.258	0.0434	0.315
p-value	(0.0003)***	(0.0005)***	(0.0171)**	(0.5472)	(0.0051)***
Sharpened q-value	[0.004]***	[0.004]***	[0.033]**	[0.460]	[0.022]**
Romano-Wolf corrected p-value	{0.012}**	{0.012}**	{0.054}*	{0.9770}	{0.018}**
<u><i>Utilitarian</i></u>	0.0607	0.0805	0.0783	0.0365	0.0345
p-value	(0.1725)	(0.1760)	(0.4280)	(0.5743)	(0.7113)
Sharpened q-value	[0.239]	[0.239]	[0.422]	[0.460]	[0.477]
Romano-Wolf corrected p-value	{0.4345}	{0.4166}	{0.8472}	{0.9880}	{0.9880}
<u><i>Malleability</i></u>	0.0897	0.102	0.155	0.0290	0.132
p-value	(0.0916)*	(0.1162)	(0.1836)	(0.6739)	(0.2758)
Sharpened q-value	[0.145]	[0.166]	[0.239]	[0.477]	[0.299]
Romano-Wolf corrected p-value	{0.1209}	{0.1529}	{0.7502}	{0.9880}	{0.3147}
Controls and School FEs	Yes	Yes	Yes	Yes	Yes
Observations	607	607	607	607	607
R-squared	0.138	0.118	0.097	0.102	0.125
P-value (BM = BMC)	0.388	0.318	0.445	0.725	0.360

Note: The figure summarizes our main results – effect of the treatments on the different gender attitude indices summarizing attitudes towards women. Standardization to mean zero and standard deviation are performed to each dependent variable. The treatments are compared relative to the placebo treated control group. The outcomes are recorded 12 months after the treatment. Women's Rights Overall is an average of all the statements concerning women's economic, social, legal, and political rights. Women's Economic Rights is an index combining women's rights to education and work outside home, based on reactions to statements "Women should be allowed to work outside the home". "Women and men should have equal rights to jobs". "I have no problem with my sister or female cousin working outside the home". "Daughters should have the same right to inherit property as sons". "Women and men should have equal rights to get an education". "Wives should not be less educated than their husbands". "Boys should not have more opportunities and resources for education than girls." Women's Political Rights is based on statements "It would be a good idea to elect a woman as the village Sarpanch (local politician)." "Women and men have equal rights to be President or Prime Minister." Women's Social Rights is based on statements "Domestic violence by husbands cannot be justified" "Parents should seek their daughter's consent before fixing her marriage". "A woman should not necessarily get married before her 25th Birthday". "Women who give birth to a son need not be honored in the family". "A woman with five daughters should not be under social pressure to bear a son." Finally, the Women's Legal Rights index is based on statements "Laws should be passed to ban dowry." "Under Article 35 of the Constitution of Pakistan & Judgment of Federal Shariat Court, the consent of 'Wali' is not required and a sui juris Muslim female can enter into a valid Nikah / Marriage under her own free will without the consent of Wali. To what extent do you approve of this legal right of women to enter marriage under their own free will." Equation (1) is estimated with all controls. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turning on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces. Further details on this are provided in Appendix D8 (considering 36 hypotheses with 4 treatments X 9 outcomes). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Impact of Teachers' Training on Petitions to Parliament

	<i>Petition to Criminalize Dowry</i>		<i>Petition to Abolish Polygamy</i>	
	(1)	(2)	(3)	(4)
<u><i>Visual narrative (movie)</i></u>	0.104	0.115	0.0598	0.0593
p-value	(0.0113)**	(0.0055)***	(0.0134)**	(0.013)**
Sharpened q-value	[0.029]**	[0.022]**	[0.029]**	[0.029]**
Romano-Wolf corrected p-value	{0.036}**	{0.023}**	{0.036}**	{0.036}**
<u><i>Joint Movie-Curriculum</i></u>	0.173	0.180	0.0866	0.0870
p-value	(0.0002)***	(0.0001)***	(0.0004)***	(0.0005)***
Sharpened q-value	[0.004]***	[0.004]***	[0.004]***	[0.004]***
Romano-Wolf corrected p-value	{0.01}***	{0.01}***	{0.013}**	{0.013}**
<u><i>Utilitarian</i></u>	0.0083	0.0070	-0.0117	-0.0106
p-value	(0.8038)	(0.832)	(0.1972)	(0.2622)
Sharpened q-value	[0.477]	[0.479]	[0.246]	[0.296]
Romano-Wolf corrected p-value	{0.9880}	{0.988}	{0.586}	{0.724}
<u><i>Malleability</i></u>	0.0089	0.0189	-0.0055	-0.0033
p-value	(0.7959)	(0.5844)	(0.567)	(0.7514)
Sharpened q-value	[0.477]	[0.46]	[0.46]	[0.477]
Romano-Wolf corrected p-value	{0.988}	{0.987}	{0.987}	{0.988}
Individual controls	No	Yes	No	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	607	607	607	607
R-Squared	0.124	0.140	0.188	0.200
Mean of dependent variable	0.114	0.114	0.030	0.030
P-value (BM = BMC)	0.185	0.206	0.433	0.409

Note: The dependent variable in Columns (1) and (2) is a dummy variable that switches on if the teacher signed a petition seeking criminalization of dowry while the dependent variable in Columns (3) and (4) is a similar dummy variable turning on for a petition seeking to abolish laws allowing polygamy in Pakistan. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turning on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces. Further details on this are provided in Appendix D8 (considering 36 hypotheses with 4 treatments X 9 outcomes). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores.*** p<0.01, ** p<0.05, * p<0.1.

Table 3: Impact of Teachers' Training on Students' Attitudes and Math Test

Scores

	<i>Student Attitudinal Survey</i>		<i>Maths</i>	
	(1)	(2)	(3)	(4)
<u><i>Visual narrative (movie)</i></u>	0.145	0.137	-0.00811	-0.00482
p-value	(0.0057)***	(0.0093)***	(0.8949)	(0.8433)
Sharpened q-value	[0.018]**	[0.021]**	[0.687]	[0.687]
Romano-Wolf corrected p-value	{0.001}***	{0.001}***	{0.992}	{0.992}
<u><i>Joint Movie-Curriculum</i></u>	0.254	0.245	0.160	0.119
p-value	p < 0.01***	p < 0.01***	p < 0.01***	p < 0.01***
Sharpened q-value	[0.001]***	[0.001]***	[0.018]***	[0.001]***
Romano-Wolf corrected p-value	{0.001}***	{0.001}***	{0.001}***	{0.001}***
<u><i>Utilitarian</i></u>	0.0713	0.0723	0.105	0.0293
p-value	(0.1767)	(0.1722)	(0.1055)	(0.2544)
Sharpened q-value	[0.245]	[0.245]	[0.178]	[0.318]
Romano-Wolf corrected p-value	{0.0709}*	{0.0599}*	{0.027}**	{0.1429}
<u><i>Malleability</i></u>	0.00171	-0.00252	0.00324	0.00419
p-value	(0.9776)	(0.9666)	(0.9584)	(0.8669)
Sharpened q-value	[0.687]	[0.687]	[0.687]	[0.687]
Romano-Wolf corrected p-value	{0.992}	{0.992}	{0.992}	{0.992}
Individual controls	No	Yes	No	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	13,911	13,911	13,911	13,911
R-squared	0.038	0.044	0.090	0.596
P-value (BM = BMC)	0.020**	0.023**	0.007***	p < 0.01***

Note: The dependent variables are standardized to mean zero and standard deviation for mathematics test scores and student attitudinal survey. The outcomes are recorded 12 months after the treatment. The corresponding survey statements from students are reported in Appendix D4. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turning on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces. Further details on this are provided in Appendix D8 (considering 16 hypotheses with 4 treatments X 4 outcomes). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Impact of Teachers' Training on Standardized Student Math Test Scores

	<i>Math Test Scores</i>				
	(1)	(2)	(3)	(4)	(5)
<i>Visual narrative (movie)</i>	-0.0384	0.0194	-0.0116	0.0106	0.0202
p-value	(0.3016)	(0.5803)	(0.7111)	(0.7533)	(0.5431)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.6084}	{0.9271}	{0.9730}	{0.9730}	{0.9041}
<i>Joint Movie-Curriculum</i>	0.164	0.0463	0.137	0.110	0.0548
p-value	p < 0.01***	(0.2258)	p < 0.01***	(0.0015)***	(0.1159)
Sharpened q-value	[0.001]***	[0.999]	[0.001]***	[0.01]***	[0.971]
Romano-Wolf corrected p-value	{0.001}***	{0.4635}	{0.001}***	{0.001}***	{0.1578}
<i>Utilitarian</i>	0.0379	0.0388	0.0256	0.0354	0.0375
p-value	(0.2782)	(0.2834)	(0.4647)	(0.3045)	(0.2972)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.5894}	{0.5894}	{0.8422}	{0.6084}	{0.6084}
<i>Malleability</i>	0.00168	-0.0204	0.0203	0.00591	0.0095
p-value	(0.9592)	(0.5847)	(0.5333)	(0.8618)	(0.7926)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.9730}	{0.9271}	{0.9041}	{0.9730}	{0.9730}
<i>Movie-Curriculum X Mixed Study</i>					0.126*** (0.0476)
<i>Mixed Study Group</i>					0.0191 (0.0310)
<i>U X Mixed Study Group</i>					-0.0147 (0.0489)
<i>M X Mixed Study Group</i>					-0.00969 (0.0475)
<i>Movie X Mixed Study Group</i>					-0.0535 (0.0487)
Students are Girls			No	Yes	
Mixed-gender Study Group Sample	Yes	No			
Individual Controls & School FE	Yes	Yes	Yes	Yes	Yes
Observations	6,959	6,952	7,107	6,804	13,911
P-value (BM = BMC)	p < 0.01***	0.485	p < 0.01***	0.005***	0.346

Note: Dependent variable is standardized to mean zero and standard deviation for math test scores. The outcomes are recorded 12 months after the treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turning on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces. Further details on this are provided in Appendix D8 (considering 20 hypotheses with 4 treatments X 5 outcomes). Mixed Study Group is a dummy that switches on when the student group is mixed-gender. *U X Mixed Study Group*, *M X Mixed Study Group*, *Movie X Mixed Study Group*, and *Joint Movie-Curriculum X Mixed Study Group* are interaction terms of Mixed Study Group with U, M, movie, and joint movie-curriculum treatments, respectively. The treatments are compared relative to the placebo treated control group. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. Table C2 reports the corresponding short-term results 6 months post-treatment. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Impact of Treatments on Social Behavior Games played against Opposite Gender

	<i>Redistribution</i>	<i>Competitiveness</i>	<i>Cooperation</i>	<i>Coordination</i>
	(1)	(2)	(3)	(4)
<u><i>Movie X Mixed Study Group</i></u>	-0.0375	-0.0666	0.171	0.184
p-value	(0.5956)	(0.3455)	p < 0.01***	p < 0.01***
Sharpened q-value	[0.999]	[0.939]	[0.001]***	[0.001]***
Romano-Wolf corrected p-value	{0.9820}	{0.8472}	{0.001}***	{0.0010}***
<u><i>Movie-Curriculum X Mixed Study Group</i></u>	-0.0406	-0.0358	0.299	0.333
p-value	(0.5457)	(0.6251)	p < 0.01***	p < 0.01***
Sharpened q-value	[0.999]	[0.999]	[0.001]***	[0.001]***
Romano-Wolf corrected p-value	{0.9820}	{0.9820}	{0.0010}***	{0.0010}***
<u><i>UX Mixed Study Group</i></u>	-0.0661	-0.0219	-0.00630	-0.0168
p-value	(0.3066)	(0.7428)	(0.8673)	(0.6270)
Sharpened q-value	[0.939]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.8052}	{0.9820}	{0.9820}	{0.9820}
<u><i>MX Mixed Study Group</i></u>	-0.0812	-0.0961	-0.0230	0.0122
p-value	(0.2064)	(0.1513)	(0.5444)	(0.7247)
Sharpened q-value	[0.703]	[0.571]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.5894}	{0.4266}	{0.9820}	{0.9820}
Playing with Opposite Gender	Yes	Yes	Yes	Yes
Individual Controls & School FE	Yes	Yes	Yes	Yes
Observations	9,145	9,145	9,145	9,145
R-squared	0.008	0.013	0.610	0.331

Note: Dependent variables are outcomes on redistribution, competitiveness, cooperation, and coordination games, respectively standardized to mean zero and standard deviation one. The outcomes are recorded 12 months after the treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces. Further details on this are provided in Appendix D8 (considering 16 hypotheses with 4 treatments X 4 outcomes). Each student plays the game twice, with the same gender and with the opposite gender. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix to:

**Transmitting Rights: Effective Cooperation, Inter-Gender
Contact, and Student Achievement**

By Sultan Mehmood, Shaheen Naseer and Daniel Chen

Contents

- A. Transcripts and IRB Discussion
- B. Additional Figures and Tables
- C. Short-Run Results
- D. Survey, Treatment Details, and Multiple Hypothesis Testing Discussion

Appendix A1. Transcripts and Discussion

Table A1: Utilitarian Treatment Transcript

I want to welcome all of you. I am your instructor for the soft skills workshop which we are starting next week. What we mean by soft skills are skills that allow us to communicate with and understand people better. The purpose of sending you a presentation is to briefly walk you through some of the core concepts which will provide you with the background knowledge that is compulsory for the upcoming workshop next week. And the first thing I want to do is, to make you comfortable. Although this is a compulsory lecture to get acquainted with the required material, there is nothing uptight about this presentation. I am really here for your benefit. I hope that is going to be a worthwhile experience for you. In this slide you see the topics that sort of headlines this presentation; We will talk about... What is empathy, Why it matters, why we need to talk about it? Then we will discuss qualitative or anecdotal evidence, some examples from teachers, to underscore the importance of empathy. After presenting evidence of these real stories from teachers, we will discuss the research which has been done on hundreds of teachers and students across many countries on empathy. Ok to begin with: In developed countries, the relevance of soft skills for student achievement in primary, middle and high school is increasingly gaining traction. More than ever before, we are talking about school culture in a way that is not primarily focused on academic achievement and passing high school. Instead, we are looking at education systems that have to come to appreciate the interdependence of academic achievement on building empathy in students. Jean Decety, a world-renowned neuroscientist, in his book ‘The Social Neuroscience of Empathy’ talks about how learning, particularly in the curriculum areas of reading, literature, and social studies, should be facilitated by empathy because

the empathic child is better able to place him or herself in the role of central characters portrayed in the fictional and historical readings. In addition to being better able to understand the roles and perspectives of these fictional and historical characters, the empathic child is better able to share and experience, to some degree, their feelings. These shared feelings may serve to underline and reinforce what they have read and been taught, resulting in better recall. Also, a number of educators have suggested that there is a reciprocal relationship between the process of reading and empathy, such that reading helps heighten and reinforce empathy (Budin, 2001; Cress & Holm, 2000). There is also evidence that teacher empathy may have a positive influence on student attitudes: teacher empathy toward withdrawn students is related with middle-school peers' acceptance of withdrawn students in their classes (Chang, 2003). This is not just limited to academic achievement though. A leading psychology and education professor at Columbia University in the US views one of the main challenges schools face today is helping students be healthy, happy and successful in meeting the challenges of their increasingly complex social environment. Empathetic teachers can play a significant role in making this happen. A research that studied teachers' behaviours in their schools to understand the role of empathy in teaching culturally diverse students in the US (McAllister, 2000), finds how teachers found it easier to work with students after being immersed into their context through interactions with their parents and their wider communities. Even the world outside school is accepting and acknowledging the importance of empathy. To contextualize the discussion with some examples, let's take the example of some of the most profitable and biggest firms across the globe. In this table you see the names of companies across the globe which scored highest points in the empathy score. That means employees and employers in these firms are rated very high in empathy. Isn't it fascinating? "It is a puzzling question for economists why the most profitable and biggest firms rank so highly in empathy scores?" Why do

firms who earn millions in profits also have high empathy scores? Is cut throatiness not going to get you more profits? Is the selfish notion of maximizing profit is most important? “Actually, it seems to be the case that soft skills are critical in all this!! “it may turn out that empathy boosts profit”. This occurs because empathy equips key partners “employees and employers with the soft skills that allow the companies to navigate complex relationships and satisfy client needs and maintain employee trust and motivation”. This empirical evidence is dispelling the view that it is being selfish and unemphatic to others is what will get you ahead in life. So, here are a few interesting definitions of empathy from different sources; this concept has been around for a while, and various religious beliefs teach us that it is something that we should practice as human beings towards others. There are different definitions of empathy in academic literature. Since there seems to be no universally agreed upon definition of empathy, “we don’t need to go into the nitty gritty of each specific definition of empathy but in a nutshell empathy is putting yourself in another's shoes”. It matters because the skill of empathy can help you succeed in your professional life. It can boost performance". That is to say, Empathy influences overall organizational performance and individual performance and well-being at a workplace. That is why, recent research is paying more and more attention to the effects of empathy on others. As we just saw in previous slide companies integrate empathy into their business strategies, because they think it’ll help them to provide better services to their clients. We don’t want to dwell too long on the private sector, but to bring it back to our context, of the importance of empathy for teachers. Empathy is important for school teachers because schools are challenging workplaces. That can be subject to emotionally demanding situations; you face demands of parents, students, school principals etc. Empathy towards yourself, toward others, and towards the students you serve can help you navigate this space better. It can help you at the job and it can improve learning for your

students, because you consciously empathize with their needs, take their point of view, understand their concerns. This is especially relevant in a country where most kids face severe hardship in daily lives and depend very much upon teachers for a safe learning environment!! For instance, think of your favorite teachers as a kid. What made them special? For me, it was their ability to understand my needs and to make me feel safe in the classroom. Teachers I could go and talk to about my concerns and I knew they would understand. And if for you that wasn't a teacher but someone else, that's okay too. The point being, we like to learn from people who can understand us. Systematic research of large population backs the idea that empathy can improve performance...also a related question is: why do all leading organizations train their employees in empathy? What is in for them? After all there is a Cutthroat competition in the corporate world for making profit. The point that I am trying to make is: in the 21st century companies might be investing in empathy to improve their profits and community engagement. A large body of research backs this up. For instance, in one prominent study at Stanford by Professor Zaki documents that empathy is more 'useful' than selfish behavior. It seems like a myth, being selfish is what will get you ahead but empathy and concern for other is key skill that those around you cherish. Empirical evidence shows that Empathy benefits at different levels. First at a personal level, empathic people report to be much happier than less empathetic people. Second at a social level, empathic people have more fulfilling social lives than less empathetic people. ” We have briefly touched on key findings from seminal studies on empathy that show empathy benefits the very people who show empathy themselves. We would have time to go into detail of these study but I did want to give you a flavor of some cutting edge research in this field...So we will go into detail of one of the studies. For instance, a research by a Yale university psychologist, Marc A. Brackett, (Brackett et al., 2011) studies 2000 students across 90 classrooms to find that there is a direct positive

relationship between classroom emotional climate, driven by teacher's ability to build strong relationships with their students. The classroom environments rated with objective indicators to be emotionally supportive had a positive impact on student conduct, suggesting that in the emotionally supportive classrooms, students liked and respected their teachers more and, in turn, behaved better. So the question is What is going on? (emphasis on it) Why is it that teachers are also investing in creating these environments? The answer many world renowned educators and psychologists (Blase, 1982; Byrne, 1994; Friedman, 1995; Hastings & Bham, 2003) say is as teachers you must have experienced the stress of when student misbehave, and this stress is not just what you feel but is proven through empirical studies showing how teacher stress, burnout and well-being have been linked consistently to student conduct. So teachers are investing their time in understanding their students and creating an environment that can support them. Empathy is social good which is valued by others. If you are empathic, your students will be more motivated to work with you FOR YOU! Empathy is mutually beneficial. Empathy helps you bring the best out of people. Only by taking the perspective of others can you realize the problems other people face in accomplishing their tasks and how they may overcome them. There are several studies that back the idea that if the teacher is empathic then the whole class performs better. Empathic leaders have better communication and trust with their peers and students. Another research on teams and performance, finds something very interesting. If you ask people on a team: who is the leader of the team? they are not likely to name the designated leader but the "effective leader who helped them out" in other words a colleague who was empathic to their needs, who may or may not be the designated leader. Again "humans are social animals", empathy begets empathy. For you teachers this is of course not a surprise. You must have heard stories of the celebrated teachers, the ones that made the difference! They incidentally also were revered not just for their work ethic and commitment to

public service but also their empathy. All types of evidence backs the idea that empathy is good for you. It is not just the right thing to do but also the most sensible thing to do for your performance as a teacher.

Table A2: Malleability Treatment Transcript

I want to welcome all of you. I am your instructor for the soft skills workshop which we are starting next week. The purpose of sending you a presentation is to briefly walk you through some of the core concepts which will provide you the background knowledge that is compulsory for the upcoming workshop next week. And the first thing I want to do is, to make you comfortable. Although, this is compulsory lecture to get acquainted with the required material but there is nothing uptight about this presentation. I am really here for your benefit. I hope that is going to be a worthwhile experience for you. In this slide you see the topics that sort of headlines this presentation. We will talk about: What is empathy? Is empathy fixed? Before going in depth in the question of whether empathy is fixed in a person. I would mention some motivating examples that point towards the notion that empathy of person is not an immutable or unchangeable force of nature. After going through the qualitative research and stories of change, I will discuss some recent large scale research that shows whether empathy changes over time? We will specifically discuss Research on Malleability of Empathy. So, here are a few interesting definitions of empathy from different sources; this concept has been around for a while, various religious beliefs teach us that it is something that we should practice as human being toward others. There are different definitions of empathy in academic literature. Since there seems to be no universally agreed upon definition of empathy, “we don’t need to go into nitty gritty of each specific definition of empathy but in a nutshell, empathy just means putting yourself in another shoes, its taking the perspective of others

when making a decision”. So the question is, Is Empathy fixed? Throughout history anecdotal accounts suggest people can change, people can change in the level of empathy they show to others (From Religion: Hazrat Umar, Khalid bin Waleed (Islamic religious leaders) and their transformation from enemy of the Islam to the greatest champions of Islam. We can find various recent examples of people who are known for their drastic transformation; growing themselves into an empathetic personality. For instance, Consider the example of Majid Nawaz from being international terrorist to running the biggest counter-terrorism organization in Pakistan (Quilliem) that fights the battle against radicalization by presenting alternative narratives to radicalized youth and actual terrorists in jails across the world (see his book “The Radical” for his fascinating story). Many other examples across the world show that people can change in the level of empathy: For instance, some White people who believed that White race is better than all in the US becoming the biggest fighters of minority rights. So, the question is what is going on? These example suggest that one can grow himself in empathy. So I made a rather bold claim based on few stories that empathy is not fixed. In fact a large body of research backs this up. For instance, in one prominent study at Stanford Zaki and co-authors show empathy is not fixed in a person. Several studies show empathy is nor fixed in a person (see e.g. Zaki and Ochsner, 2012). “Empathy is changeable and can be influenced over time. Empathy is not stable over one’s lifetime. It can be developed and cultivated.” Survey after surveys also show that empathy of populations changes over time. An important point is: Empathy doesn’t come naturally in all situation: For instance: Sometimes we struggle with showing empathy for someone or considering their perspective. That’s OK, empathy can be changed. If we don’t feel empathy naturally, it doesn’t mean that we are incapable of feeling it. empathy is changeable, and that understanding that it

can sometimes be difficult to feel empathetic unless we work on it: is important step to developing this important life skill. Another important point is “Empathy is not a constant of nature determined by your upbringing alone, it rises and falls based on the environment around you”). For instance, in United States where most amount of data is available empathy scores have been falling for the last 30 years i.e. empathy in US now is about 50% of what it was 30 years ago. Why is it falling, if it is fixed? And it is not just one measure of empathy but all measures seem to follow this downward trend. This data convinced many psychologists that empathy is malleable, people can grow in empathy or they can fall in empathy. That is exactly what this graph indicates: that empathy is falling over time! If empathy is fixed theory is correct, this graph would not be downward trending. It should be a straight line. Essentially, this is inconsistent with the fixed empathy theory where empathy of individual and populations are fixed over time. This observed decline has put out of business all the psychological theories that had argued earlier empathy was fixed. We have briefly touched on key findings from the seminal study on empathy that show empathy is not fixed. I do want to give you some more flavor of cutting-edge research in this field. So we will go into detail of couple of the studies. For instance: does empathy change? empathy changed when they were given perspective of others (VR glasses, research article: Bernard et al., 2018). In the first study when researchers gave virtual reality goggles to people and made them take perspective of others (e.g. see the lives through the eyes of homeless people and beggars), the level of empathy they showed to others skyrocketed both in surveys as well as high-stake decisions). Therefore, being open minded and willing to change and learn, is essential to grow in empathy and develop this skill. A seminal study from Stanford University shows that people who are most rigid in their believe that empathy cannot change in them or others are the least empathetic

to begin with. People who believe empathy is inherent and unchangeable disengage from situations where empathy is difficult for them to experience. By contrast, people who believe empathy can be developed, they feel less threatened by perceiving that their empathic abilities are being challenged in a difficult situation. Another study shows that “Resilience training” increased empathy among radicalized Moroccan youth (research article: Feddes et al., 2015). This suggests that “People really change? it hints towards the notion that we need to revise this notion empathy cannot be changed and is fixed, the level of empathy an individual has is not destiny. This also suggests the answer of the puzzling question why the most profitable and biggest firms engage in empathy workshops and “waste” millions if empathy is unchangeable? Can it be that companies like Google and Facebook think empathy is malleable in people? They can inculcate these skills So, coming back to the basic question we began with, can empathy evolve in a person? Commonsense stories, all types of evidence point to one conclusion that empathy is malleable and it can change. Empathy is a skill that can be developed. Like any skill it needs work, to understand the needs of others and not just to best serve them but bring the best out of your students. Learning “The art of empathy” needs practice. All types of evidence backs the idea that empathy is not fixed but malleable. It is a skill that can be developed.

Table A3: Transcript of Structured Discussion

Each training workshop was followed by a 30-minute structured group discussion within the treatment arm which was structured as follows.

2 candidates were randomly drawn from each treatment arm and were invited to speak and answer within the Utilitarian, Malleability, Bol Movie, the Bol Movie joint with Curriculum and Placebo arms, these two questions:

Randomly Drawn Candidate 1 within the Treatment Arm:

Q1. What do you think were the main messages of this workshop? Q2. How do you think you may apply lessons from today's workshop in your career?

Randomly Candidate 2 within the Treatment Arm:

Q1. What struck you most about today's workshop and why? Q2. Can you give examples on how the lessons of today's workshop could be applied in your teaching job?

Table A4: Research Ethics Approvals

Research Ethics Approvals.— Our study protocols were reviewed and approved by the Lahore School of Economics with IRB Number RERC-062021-03. Earlier, we had also received separate administrative approvals from the Progressive Education Network administration, and teacher representatives. We also closely collaborated with the PEN training department to design and implement the experiment, who shared our view that this project is important and may bring out important lessons for policymakers and for their own inhouse training programs,

allowing them to improve both teacher satisfaction and student achievement in the future.

Table A5. Discussion of Curriculum Treatment

The "*Gender Equality in Child Development for Social Skills*" course is a collaborative endeavor crafted by a team of local educators, our research team, and the research staff at Progressive Education Network. Leading this initiative are Miss Maeeha Farooqi and Miss Sameen Amir, two esteemed public school educators affiliated with the Progressive Education Network in Pakistan, who have two decades of experience in the teaching profession. This curriculum has a flexible structure that allows educators to conduct sessions at a pace of one hour per week or adjust the frequency to accommodate the school's schedule. The overarching goal of this curriculum was to discuss gender rights, critical thinking and discussion on the respect for gender differences. The curriculum also attempts to foster a culture of trust and cooperation, and empowering students to challenge gender stereotypes and discrimination. The weekly curriculum is a collection of interactive and reflective activities, each with a specific purpose and outcome aimed at nurturing emotional intelligence on the topic of gender rights. The program begins with an introduction to the concept of gender, challenging preconceived biases and encouraging respect for individual expression. Subsequent weeks delve into recognizing and understanding emotional signs across genders, the universality of emotions, and the importance of gender sensitivity and rights. Activities range from role-playing and storytelling to creative arts, all designed to facilitate a deeper understanding of gender dynamics and emotional experiences. The curriculum also includes practical exercises such as the "Emotion Maze," which teaches perspective-taking, and decision-making scenarios that promote inclusivity and respect for gender diversity. These activities are complemented by case scenarios that reflect societal gender discrimination, prompting students to engage in critical discussions and self-reflection on gender attitudes. The full coursebook this course is based on can be access in full [HERE](#).

Table A6. Discussion of Teacher Workshop

A structured three-hour workshop was conducted for teachers from the Progressive Education Network (PEN) schools, focusing on the "Gender Equality in Child Development for Social Skills" course in Lahore, Garden Town PEN Headquarters. The curriculum workshop took place in March 2021, followed by the launch of the course in April 2021, which continued for a duration of four months. This gender course was incorporated into the existing Pakistan studies class schedule. The workshop's agenda, followed the coursebook we adapted from Professor Sule Alan, it encompassed a review of the curriculum's adaptable framework, which permits educators to modulate the frequency of sessions to align with institutional schedules of their specific school. Emphasis was placed on the curriculum's primary objectives: to facilitate discourse on gender rights, to cultivate critical thinking, and to instigate conversations on the acknowledgment of gender differences. The workshop equipped teachers with pedagogical strategies to engender a classroom environment conducive to trust and collaborative learning, while also empowering students to critically assess and challenge entrenched gender stereotypes and biases. The curriculum's array of interactive and reflective exercises, including role-playing, storytelling, and creative arts, was delineated, each designed to enhance students' emotional intelligence on gender rights. The workshop also provided a platform for teachers to engage with case scenarios reflective of societal gender discrimination, thereby encouraging critical discussions and introspection on gender attitudes among students. Ultimately, the workshop for PEN teachers at Lahore's PEN Headquarters equipped teachers with the skills to deliver a flexible, gender-focused curriculum that aims to shape a more inclusive and empathetic

classroom environment, using the course book we designed to deliver the content. The full coursebook we followed in the workshop can be accessed [HERE](#).

Appendix A2. Robustness and Discussion

Experimental Demand.— Experimental demand is unlikely to drive our results for at least four reasons. First, we observe a virtually zero effect on teachers and students of our utility and malleability treatments, which attempt to impact teacher gender attitudes through interventions building on evidence from the economic and psychology literature. Second, we use the implicit association test, which is difficult to “game” and captures implicit associations that may be unknown even to the individual (Greenwald et al., 2009). Third, the persistence of the effects of our visual narrative alone and joint narrative-and-curriculum treatments up to a year later, particularly on student achievement, strengthens the inference that the treatments had real consequences. Last, we use the Marlowe-Crowne social desirability scale, a survey module developed by social psychologists to rigorously measure a person’s propensity to give socially-desirable answers (Crowne and Marlowe, 1960). In Table B6 of Appendix B, when we discard individuals who score high on this social desirability scale, the results remain essentially identical. Taken together, all this indicates that our attitudinal data is likely to map onto real-world behavior.

Spillovers.—Our experiment allowed us to randomly allocate treatment for 607 teachers across 52 schools and 15000 students in Pakistan. However, students and teachers in the treated and control group may interact within a school. This can lead to potential spillovers if individuals in the control group also end up being partially treated. First, to the extent there are spillovers within a school, our estimate can then be considered as a lower bound on the impact of the treatments. Second, even if the spillover goes into the opposite direction, their effect is likely

to be limited. This is because our experimental design allows us to ascertain the extent of these spillover effects. That is, we exploit the variation in treated teachers across the 52 PEN schools in our sample to explore how it impacts teacher gender attitudes and student gender attitudes and test scores. Panel A of Table B7 shows that teachers' gender attitudes are essentially identical as more teachers get treated within a school. Likewise, in Panel B of Table B7, we find that the treatment effect on student's gender attitudes and test scores are also similar as more teachers within a school get treated. In summary, regardless of the fraction of treated teachers in a school, the visual narrative and joint treatment has similar impacts. Last, we leverage the fact that our Utilitarian and Malleability treatments had no impact on gender attitudes over the control group. Therefore, we investigate whether the fraction of schools treated with visual narrative or joint treatment causes the control group to increase in progressive gender attitudes. Under the assumption that a higher fraction of treated teachers leads to a greater likelihood for interactions between treated teachers and control teachers, we assess the impact of fraction of treated teachers on the control teacher's outcomes. These results are reported in Tables B8 and B9, where we find little impact of fraction of treated teachers among control teachers on gender attitudes. The null effect of more intensely treated schools on gender attitudes holds for Utilitarian, Malleability and Placebo assigned teachers. Taken together, the evidence strongly suggests that spillover effects between treated and control teachers, even if they exist, are likely to be small in magnitude and hence unlikely to explain the effect sizes we observe.

Attitudes to Behavior.— Four aspects of our experiment offer interpretation of attitudinal data as having real consequences. First, we use an implicit association test, which is based on the idea that the easier the mental task, the faster the response. IATs have the advantage of (1) mitigating

social-desirability bias in the responses and (2) capturing implicit associations that may be unknown to the individual but may nevertheless affect attitudes and behavior (Greenwald et al., 2009). Second, we observe that the students' math scores are impacted, and students' gender attitudes shifted even when the standalone visual narrative treatment focused squarely on teachers. Finally, following Dhar et al., (2022) we use the Marlowe-Crowne social desirability scale, a survey module developed by social psychologists to rigorously measure a person's propensity to give socially-desirable answers (Crowne and Marlowe, 1960). When we discard individuals who score high on their social desirability scale, the results are essentially identical. These features of our experiment indicate that our attitudinal data is likely to map well onto real-world behavior.

External Validity.—Although comprehensive data on charter schools in Pakistan is scarce, by the early 2020s, a number of organizations inspired by the U.S. charter school model had begun to effectively compete with both public and private schools. This marked a significant shift from virtually no charter schools in the early 2000s to over 3,000 operating across Pakistan. For example, our network of charter schools, the Progressive Education Network which started with only 52 schools in Punjab, where we intervened, at the beginning of our experiment in 2020, has expanded to 231 schools in Punjab by 2024. This exponential growth has been driven in part by substantial foreign donations, especially from Pakistani expatriates in the United States who wished the 'replicate' the US charter school experience in Pakistan (Isensee, 2016). Our study was conducted in Punjab, the most populous province in Pakistan with a population of 110 million. The province is home to approximately 100,000 public schools, 30,000 private schools, and 3000 charter schools (National Center for Education Statistics, 2019). Since the early 2010s, Pakistan has witnessed a significant surge in the number of private and charter schools. For example, the number of private schools has doubled since 2000, reaching approximately 60,000

by 2016 (Andrabi et al., 2024). Even though charter schools may not be the most widespread type of school in Pakistan, their growth in adopting public schools is noteworthy. With the rapid proliferation of charter schools in Pakistan, it is becoming increasingly important to study their impact and development within the country's educational system. This research could also provide insights into how the U.S. charter school experiment might be 'externally valid' in other contexts.

Appendix B: Additional Figures and Tables


Figure B1: Summary of All Studies with the Progressive Education Network in Punjab and Sindh Province

All Experiments with Teachers and Students of the Progressive Education Network (PEN)					
Title of Paper	Sample	Timing of Experiment and Treatment Rollout	Main Outcomes Observed	Treatments	Results
Transmitting Rights	607 teachers and their 13,932 pupils in 52 PEN schools in Punjab, Pakistan.	The Baseline, midline, and endline surveys were conducted in February 2021, September 2021, and March 2022, respectively. We conducted the first round of data collection 6 months after the intervention (labelled short-term effects), while the second round of data was collected 12 months later. The treatment of Transmitting Rights and Rights Revolution happened at the same time but gender attitude outcomes for Transmitting Rights were collected earlier (February 2021, September 2021, and March 2022). The stress outcomes were collected later on request of PEN administration and were not pre-registered. Note all treatments were rolled out in March 2021.	Teachers' gender attitudes, IAT scores, Students' Attitudes and Math Test Scores	Utilitarian Treatment, Malleability Treatment, Visual Narrative Treatment, Joint Visual Narrative and Gender-Rights Curriculum Treatment. Different practices of shifting the gender	Visual narratives and joint visual narratives combined with the curriculum foster more progressive gender attitudes among teachers. These gender attitudes are transmitted from teachers to students. Broad-based interventions, such as those based on Utilitarianism and the Malleability of Empathy, have no effect.
Why Are Rights Revolutions Rare?			Stress Outcomes, Domestic Violence		A year after the intervention, the cultivation of progressive gender attitudes elevated stress hormone concentrations in the blood plasma. However, this reduces as more teachers within a school are treated and have more progressive gender attitudes, a moral bandwagoning effect.
Role Models and Theory of Mind: Teacher Vaccinations and Student Success			All treatments was rolled out in August 2021, with the baseline collected 6 month before treatment (February 2021), midline 12 months (September 2022) post-treatment and endline 18 months post-treatment (March 2023), respectively. For mathematics, we have test scores for 6, 12 and 18 months after the treatment, and vaccinations and absenteeism data is available at the monthly level up to 18 months post-treatment.	COVID-19 Vaccination Certificates, Mathematics Test Scores	Cash 15%, Cash 30% and Lottery Treatment, Celebrity Treatment, Role Model Treatment. Methods of encouraging people to get vaccinated
Psychological Well-being and Educational Efficiency: Evidence from Civil Servant Intervention in Pakistan	850 teachers and their 24,752 students of PEN in province of Sindh and Punjab.	All treatments were rolled out in February 2022 (ongoing study). To monitor student performance, we utilized the February 2022 scores as the baseline, representing pre-intervention levels, and the September 2022 scores as the endline for assessing post-intervention outcomes.	Stress, Test Scores, Teacher Attendance.	Cognitive Behavioral Therapy, Psychopharmacological Therapy, Mindfulness Meditation	Mindfulness Meditation (MM) and CBT effectively reduce stress, with psychopharmacological aid showing less impact.
Reshaping beliefs about ourselves and others: Experimental Evidence from civil servants in Pakistan	400 teachers and their 10,000 students of PEN in province of Sindh.	All treatments was rolled out in January 2022. The baseline survey was carried out in the second week of January, the midline survey in August, and the endline on 17th December 2022.	Growth Mindset Scores, Student Test Scores.	Growth Mindset Treatment, Narrative Treatment, Empirical Evidence Treatment	Growth mindset training shifts teachers' beliefs about malleability of intelligence, and reduces stereotypes against first-generation learners and students from disadvantaged backgrounds. In contrast, exposure to narrative or empirical evidence about teacher value-added did not have statistically significant effects.


Note: The table presents a synthesis of five distinct studies conducted in collaboration with the Private Education Network (PEN) in Pakistan during the 2020s. The first three studies listed were carried out in the same Punjab PEN schools, while the last two studies also included PEN schools from Sindh province. The studies cover various samples, timelines, outcomes, and treatments, with details provided within the table. It is important to note that the samples across these studies are not identical, except for the first three papers, which utilized the same sample of 52 PEN schools and 607 teachers in Pakistan's Punjab Province.

Figure B2: Lesson Plans of Teachers

Panel A: Lesson Plan of Teacher A in English

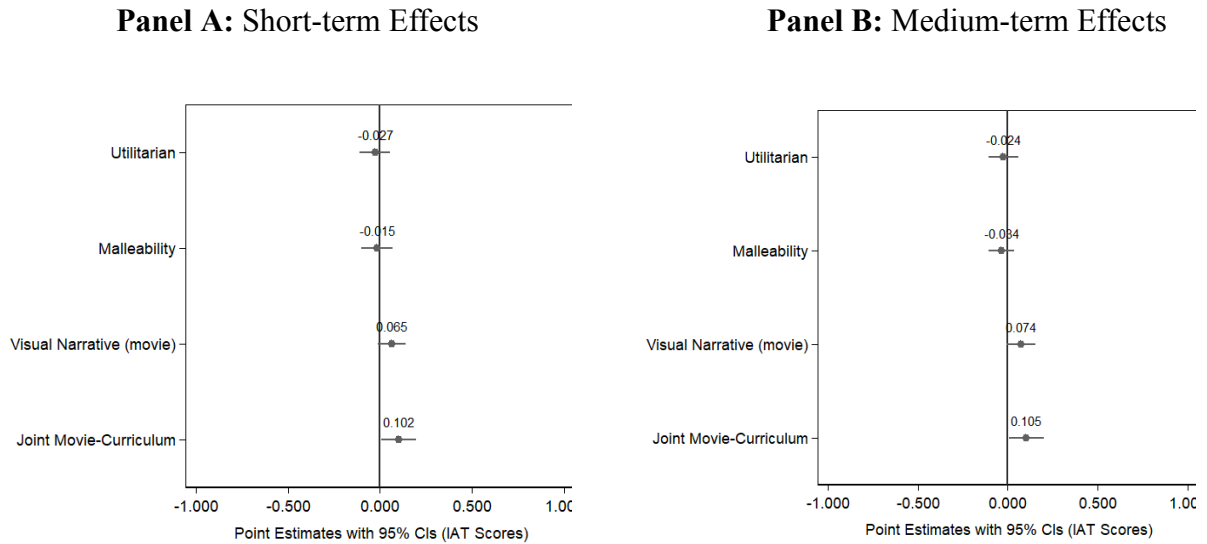
Subject/Date	Topic/Page#	No. of Periods	Pedagogical Strategy/Teaching Methodology	Resource Material (AV Aids)	Assign-ment
			Q → Would you want the job of your mother or your father, when you grow up? Part 1 :- Ask the class on what their mothers' & fathers' do. Part 2 :- Ask the class, what they would choose to become, from what their parents do?	- Book - Pictures - Board.	Write a note on what your job in future is more likely to look like? And does it resemble any of your parent's job?
			Part 3 :- Show them the following pictures. 		
			Part 4 :- Initiate an open discussion on what the children choose from the above pictures as their job & why?		

Panel B: Lesson Plan of Teacher B in Urdu

Subject/Date	Topic/Page#	No. of Periods	Pedagogical Strategy/Teaching Methodology	Resource Material (AV Aids)	Assign-ment
			کیا آپ اپنی امی اور ابو کی نوکری کو بڑے بڑے کاموں پر چاہیں گے۔ حصہ ۱ :- جماعت سے پوچھیں کہ آپ کے امی اور کیا کرتے ہیں حصہ ۲ :- جماعت سے پوچھیں کہ وہ اپنی امی ابو کی ملازمتوں میں سے کس ملازمت کا انتخاب کریں گے۔ حصہ ۳ :- بچوں کو تصاویر دکھائیں۔	کتابیں تصاویر بورڈ	مشتمل ہیں آپ اپنے امی ابو کی کس کی ملازمت کو چاہیں اور کیوں
					
			حصہ ۴ :- بحث دہرائیے قائم کریں اور ہر دی گئی تصاویر پر		

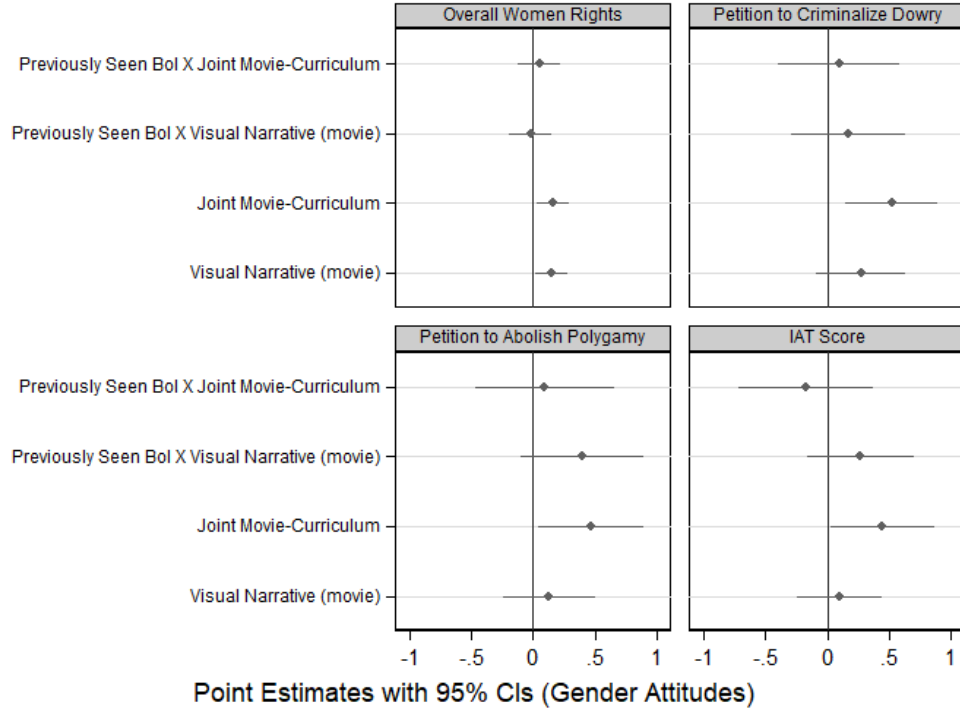
Note: The teachers entered the gender rights curriculum in their daily lesson plan logs and taught in the language they were most comfortable with. Teacher in Panel A logged her teaching activities in English, while the teacher in Panel B logged it in Urdu.

Figure B3: Impact of Utilitarian, Malleability, Bol Movie, and Bol Movie and Curriculum Treatments on Original IAT Scale



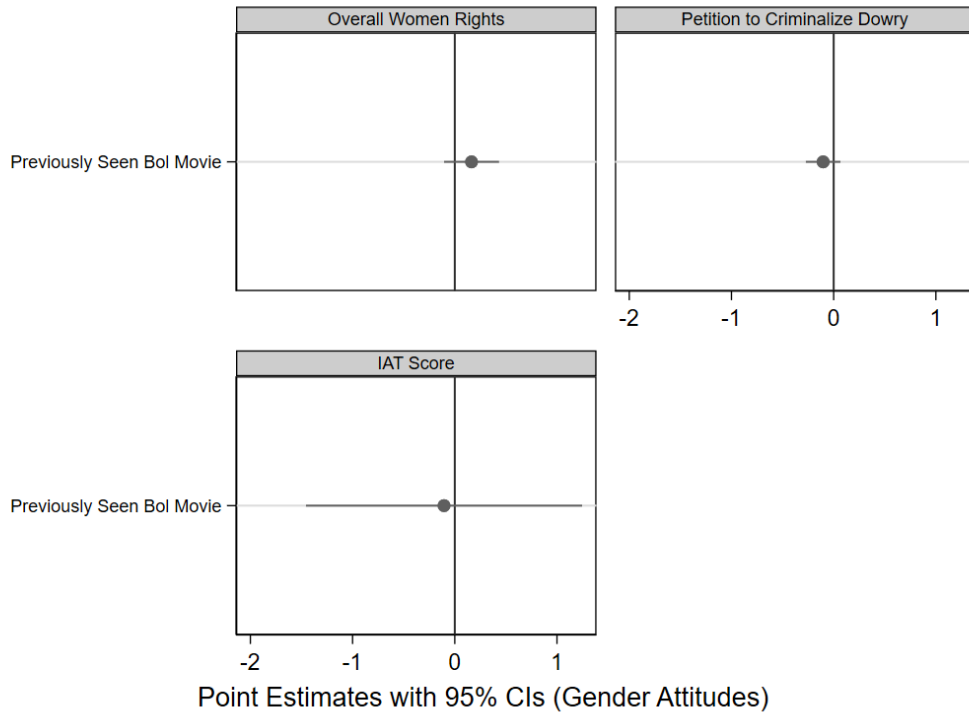
Note: The figure reports estimates from equation (1) with all controls. In Panel A, the raw IAT (Implicit Association Test) score in original units is used as a dependent variable with the effect shown in the short-term (6 months). The IAT value ranges from -1 to 1 (Greenwald et al., 2009). If the magnitude is less than 0.15 (for both positive & negative scores), the score shows, there is little to no association. If the magnitude is between 0.15 and 0.35 (for both positive & negative scores), the score shows, there is slight association. If the magnitude is between 0.35 and 0.65 (for both positive & negative scores), the score shows, there is moderate association. If the magnitude is greater than 0.65 (for both positive & negative scores), the score shows, there is strong association. In Panel B, medium-term or a year after treatment effects are shown. Controls include all individual characteristics and school fixed effects. 95% confidence bands are also reported.

Figure B4: Impact of Previously Seeing the Bol Movie



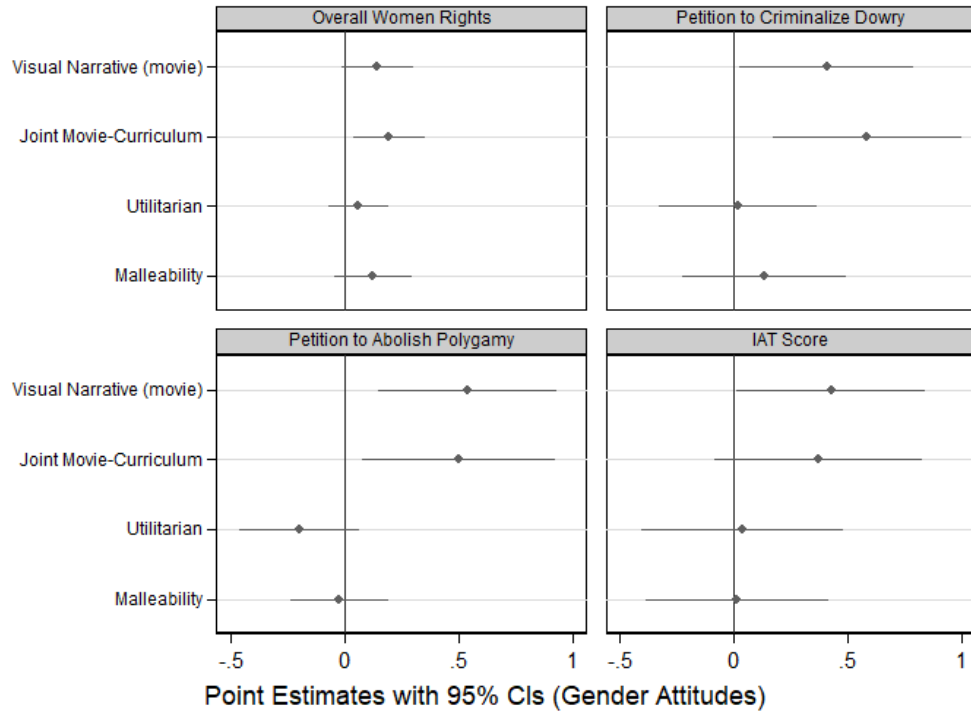
Note: The figure for Panel A reports estimates from equation (1) with all controls and school fixed effects with Gender Rights Index, Petitions and IAT as the dependent variables in each panel above. The dependent variables are standardized to mean zero and standard deviation one. The *Previously Seen Bol* is a dummy variable that switches on if the teacher is reported to have previously seen the Bol movie. *Visual narrative (movie)* represents the visual narrative treatment of the movie Bol augmented with the 30-minute structured discussion of gender rights themes in the movie. *Joint Movie-Curriculum* is the dummy that switches one for participants who received the visual narrative of the movie (along with the structured discussion) together with the gender studies curriculum. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. 95% Confidence Intervals are also reported.

Figure B5: Gender Attitudes, Petitions and IAT and Previously Seen Bol Movie for the Control Group Subsample



Note: The figure report estimates from regression for Gender Rights Index, Petitions and IAT as the dependent variables in each panel, respectively. The dependent variables are standardized to mean zero and standard deviation one. The Previously Seen Bol is a dummy variable that switches on if the teacher is reported to have previously seen the Bol movie. No one in the control group who had previously watched the "Bol" movie signed the petition against polygamy. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. 95% Confidence Intervals are also reported.

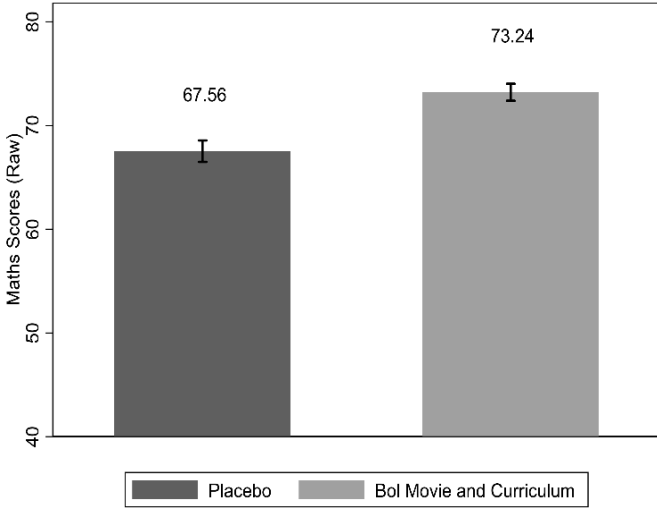
Figure B6: Effect of Visual Narrative and Joint Treatment on the Teachers for the Ever Watched Bol Subsample of Teachers



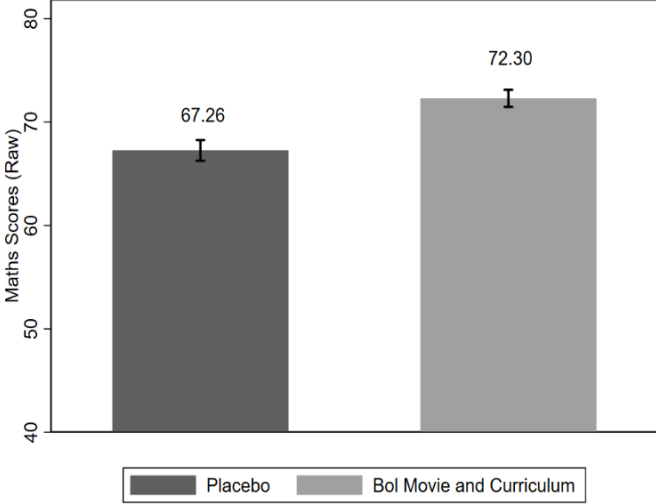
Note: The figure report estimates from equation (1) for people who have previously watched ‘Bol’ movie with all controls. The dependent variables are Gender Rights Index, Petitions to parliament and gender IAT test. The dependent variables are standardized to mean zero and standard deviation one. Visual narrative (movie) represents the movie Bol augmented with the 30-minute structured discussion of gender rights themes in the movie. Joint Movie-Curriculum is the dummy that switches one for participants who received the visual narrative of the movie (along with the structured discussion) together with the gender studies curriculum. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. 95% Confidence Intervals are also reported.

Figure B7: Difference in Math Scores following Joint Movie-Curriculum and Placebo treatments by Type of Study Group
Mixed-Gender Study Group

Panel A: Short-term Effects

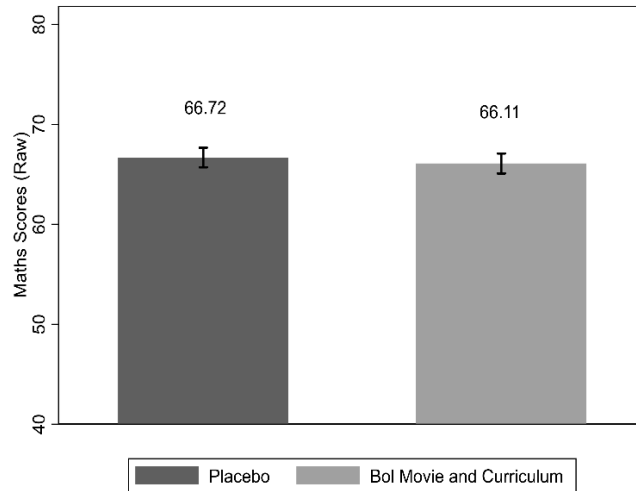


Panel B: Medium-term Effects

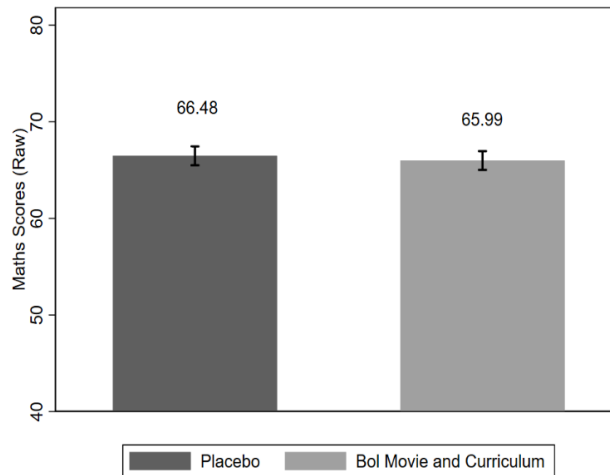


Same-Gender Study Group

Panel C: Short-term Effects



Panel D: Medium-term Effects



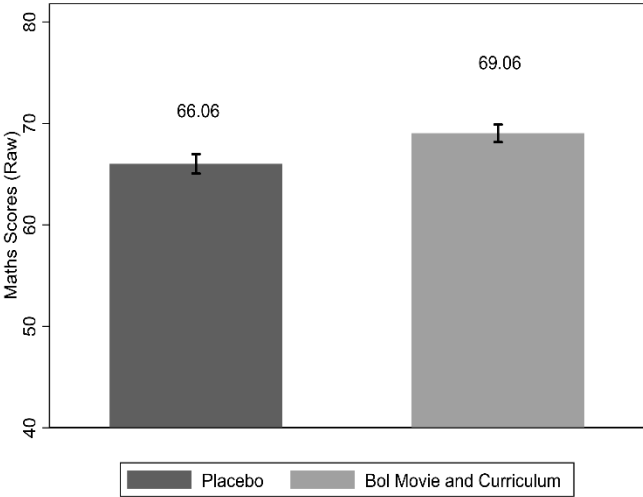
Note: Short-term and medium-term mathematics scores are shown on the right and left, respectively. Short-term effects in Panel A and C are recorded 6 months after the treatment and medium-term effects in Panel B and D are recorded twelve months after the treatment. Scores range from 0 to 100. The upper panels (Panel A and B) display scores for mixed-gender study groups, the lower panels (Panel C and D) for same-gender study groups, comparing students' average scores under the joint movie-curriculum treatment with the placebo group. 95% confidence bands are also reported.

Figure B8: Difference in Math Scores following Joint Movie-Curriculum and Placebo treatments

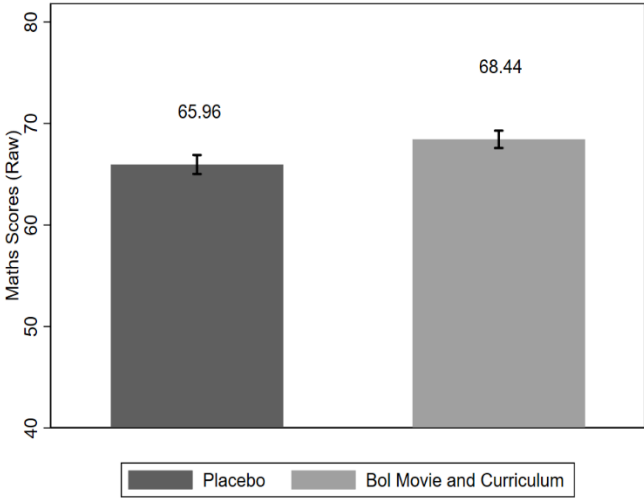
by Gender

Boys' Mathematics Scores

Panel A: Short-term Effects

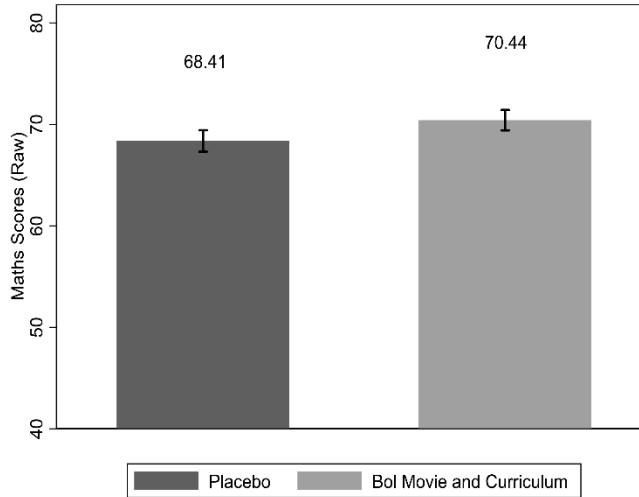


Panel B: Medium-term Effects

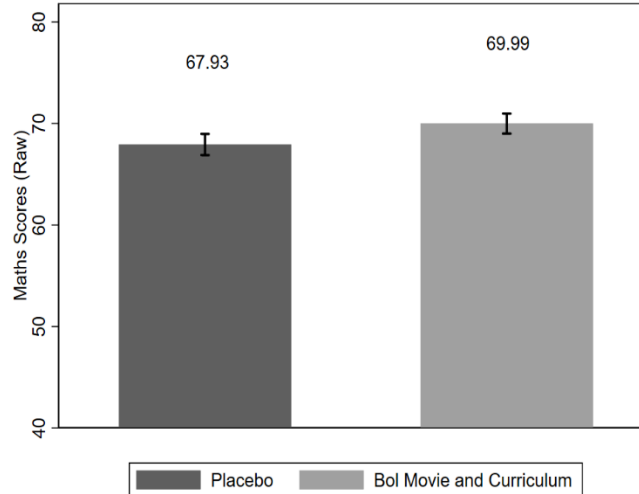


Girls' Mathematics Scores

Panel C: Short-term Effects



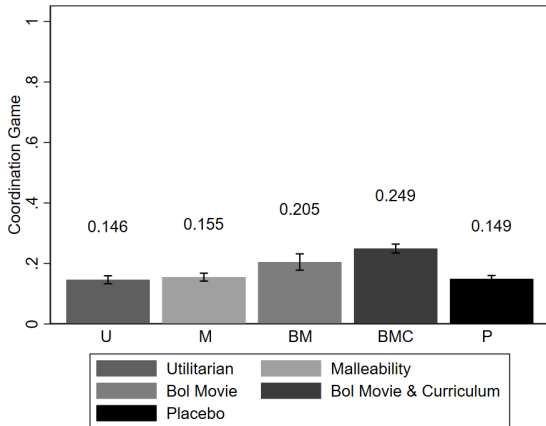
Panel D: Medium-term Effects



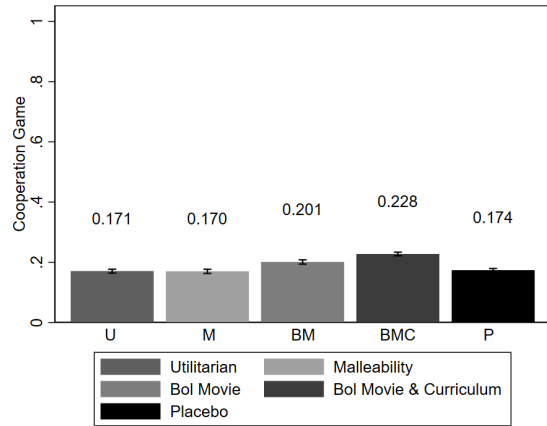
Note: Short-term and medium-term mathematics scores are shown on right and left, respectively. Short-term effects in Panel A and C are recorded 6 months after the treatment and medium-term effects in Panel B and D are recorded twelve months after the treatment. Scores range from 0 to 100. The upper panels (Panel A and B) display math scores for boys, the lower panels for girls.

Figure B9: Game fielded with the Opposite Gender

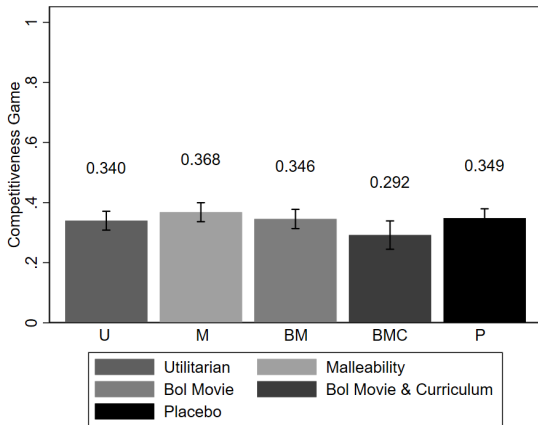
Panel A: Coordination



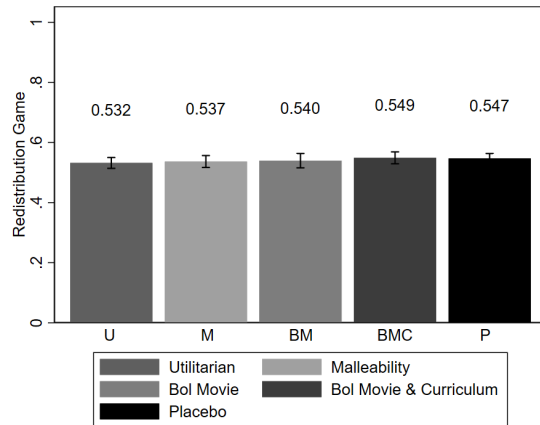
Panel B: Cooperation



Panel C: Competitiveness



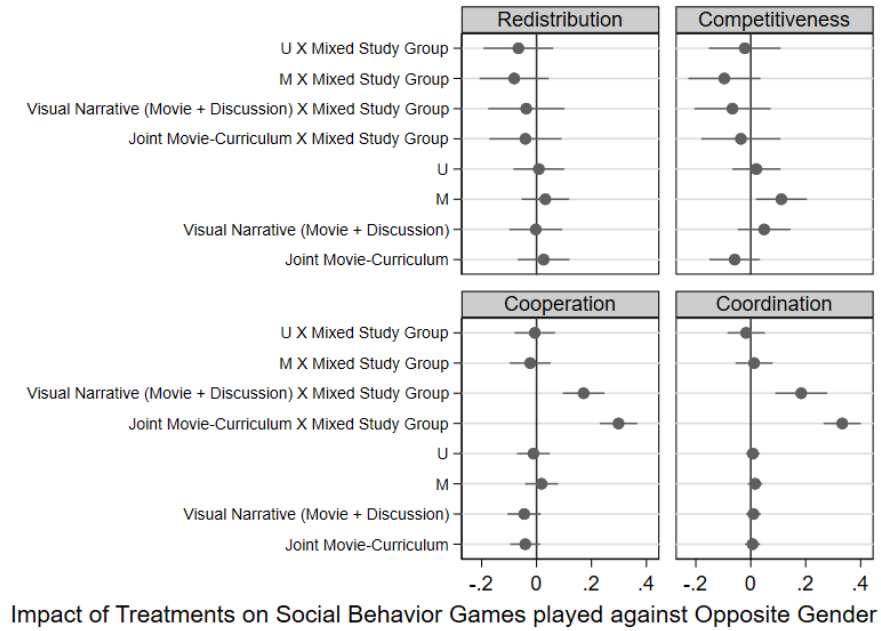
Panel D: Redistribution



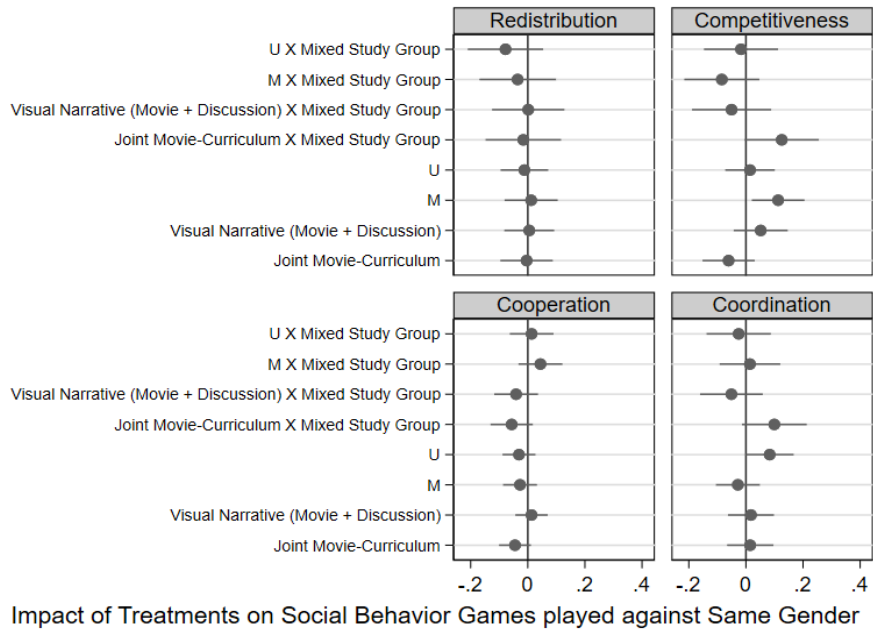
Note: Figure above displays averages on outcomes on redistribution, competitiveness, cooperation, and coordination games in their respective marked panels when the student is playing against the opposite gender. The figures show the level effects of all treatments.

Figure B10: Impact of Treatments on Social Behavior Games played against Opposite vs Same Gender

Panel A: Opposite Gender



Panel B: Same Gender



Note: Figure above displays outcomes on redistribution, competitiveness, cooperation, and coordination games in their respective marked panels when the student is playing against the opposite gender. The figures show the level effects of all treatments as well as interactive effects with mixed (Panel A) versus same-gender (Panel B) study group.

Table B1.1: Summary Statistics

Panel A: Teacher Data					
	Obs	Mean	Std. Dev.	Min	Max
Overall Women's Rights (short-term)	607	1.537	.363	1	3.563
IAT Score (short-term)	495	.051	.308	-.928	1.116
Overall Women's Rights (medium-term)	607	1.628	.368	1	3.750
Petition Against Dowry	607	.114	.318	0	1
Petition Against Polygamy	607	.03	.17	0	1
IAT Score (medium-term)	527	.062	.301	-.999	1.124
Teaching Experience	607	4.608	2.957	0	18
Years of Education	607	12.679	1.511	10	18
Professional Qualification	607	.259	.438	0	1
Panel B: Student Data					
	Obs	Mean	Std. Dev.	Min	Max
Gender	13,911	.511	0.500	0	1
Kindergarten	13,911	.169	.375	0	1
Grade 1	13,911	.163	.369	0	1
Grade 2	13,911	.173	.379	0	1
Grade 3	13,911	.153	.360	0	1
Grade 4	13,911	.144	.351	0	1
Grade 5	13,911	.131	.337	0	1
Pre-Treatment Gender Index	13,911	77.703	18.334	5	100
Pre-Treatment Math Score	13,911	66.269	20.451	1	99
Students' Gender Attitudes	13,911	82.566	18.881	3	100.4
Students' Math Scores	13,911	67.416	18.824	-1	102

Note: Overall Women's Rights is an index consisting of average of all 16 statements concerning Women's Economic, Social, Legal, and Political Rights listed in Appendix D. Women's Economic Rights is an index combining women's rights to education and work outside the home, based on reactions to statements "Women should be allowed to work outside the home". "Women and men should have equal rights to jobs". "I have no problem with my sister or female cousin working outside the home". "Daughters should have the same right to inherit property as sons". "Women and men should have equal rights to get an education". "Wives should not be less educated than their husbands". "Boys should not have more opportunities and resources for education than girls." The IAT Score is the gender Implicit Association Test (IAT). Petition Against Dowry is a dummy variable that switches on if the teacher signed a petition sent to parliament seeking criminalization of dowry, while the variable Petition Against Polygamy is a petition to abolish laws allowing polygamy in Pakistan. In Gender, Male notes as 1 and Female as 0. Variables associated with Grade are dummy variables that switch on if a student studies in the corresponding grade. The pre-treatment Gender Index is an index concerning the Gender Attitudes of students before the treatment. Pre-treatment Math Score is the math scores of students before treatment. The total number of available observations, means and standard deviations of variables are also reported.

Table B1.2: Balance over Teacher Characteristics and Pre-Treatment Gender Attitudes

	Pre-Treatment Gender Index	Movie “Bol” Watched	Married	Av. Teaching Hours	Av. Class Size	Teaching Experience	Years of Education	Educational Specialization
<i>Visual narrative (movie)</i>	-0.0720	0.0869	-0.0767	0.418	-2.323	0.0238	0.0719	0.0109
	[0.0775]	[0.0675]	[0.0674]	[0.356]	[2.708]	[0.400]	[0.206]	[0.0589]
<i>Joint Movie-Curriculum</i>	0.00463	0.0598	-0.0619	0.130	-1.409	0.481	0.138	-0.0125
	[0.0789]	[0.0688]	[0.0687]	[0.362]	[2.757]	[0.407]	[0.210]	[0.0600]
<i>Utilitarian</i>	-0.121	0.0265	0.0107	-0.0927	-0.666	-0.0173	-0.112	-0.0539
	[0.0768]	[0.0669]	[0.0668]	[0.353]	[2.683]	[0.396]	[0.204]	[0.0584]
<i>Malleability</i>	-0.0429	0.0766	-0.0965	0.132	-2.167	0.441	-0.0771	-0.0559
	[0.0784]	[0.0683]	[0.0682]	[0.360]	[2.737]	[0.404]	[0.209]	[0.0596]
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	607	607	607	607	607	607	607	607
R-squared	0.105	0.075	0.078	0.063	0.088	0.075	0.057	0.085
Mean of dependent var	-0.071	0.532	0.473	30.28	24.81	4.608	12.679	0.259
P-value (BM = BMC)	0.329	0.692	0.829	0.424	0.739	0.259	0.752	0.694

Note: Robust standard errors appear in brackets (clustered at teacher level). The treatments are compared relative to the placebo treated control group. Dependent variables are teacher-level individual characteristics. Teaching Experience is years of teaching experience. Years of Education is the number of years of formal studies. Educational Specialization is a dummy variable that switches on for a specialization in Pedagogy. Av. Teaching Hours is the average number of hours taught every week. Av. Class Size is the average number of students in the class. Married is a dummy variable that switches on if a teacher is married. In Panel B, the dependent variables are available student-level individual characteristics and pre-treatment math scores (in the previous year endline tests). Gender is a dummy variable that switches on for a female student. Grades 1 to 6 are dummy variables that switch on for a student from grades one to six, respectively. Pre-Treatment Math Scores are standardized math scores held a year before our treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B1.3: Balance over Student Characteristics and Pre-Treatment Gender Attitudes

	Gender	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Pre-Treatment Gender Index	Pre-Treatment Math Score
<i>Visual narrative (movie)</i>	-0.00445 [0.0243]	0.0242 [0.0454]	0.0863* [0.0480]	-0.0392 [0.0464]	0.0162 [0.0442]	-0.0448 (0.0358)	0.00101 (0.0267)	-0.0202 (0.0552)
<i>Joint Movie-Curriculum</i>	0.00169 [0.0272]	-0.0483 [0.0443]	0.0598 [0.0481]	-0.0690 [0.0455]	0.00437 [0.0414]	0.0846* (0.0450)	-0.0360 (0.0271)	-0.0469 (0.0618)
<i>Utilitarian</i>	-0.00107 [0.0274]	0.0814* [0.0490]	-0.0348 [0.0446]	-0.0456 [0.0473]	-0.0323 [0.0419]	0.0579 (0.0412)	-0.0272 (0.0267)	0.0770 (0.0592)
<i>Malleability</i>	0.0168 [0.0277]	-0.00418 [0.0467]	0.0208 [0.0458]	-0.0111 [0.0491]	-0.0235 [0.0413]	0.0385 (0.0442)	0.00069 (0.0266)	0.0166 (0.0599)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,932	13,932	13,932	13,932	13,932	13,932	13,932	13,932
R-squared	0.360	0.046	0.034	0.019	0.039	0.039	0.005	0.098
Mean of dependent var	0.511	0.163	0.173	0.153	0.144	0.131	0.000	0.000
P-value (BM = BMC)	0.795	0.099*	0.604	0.489	0.794	0.002***	0.185	0.661

Note: Robust standard errors appear in brackets (clustered at teacher level). The treatments are compared relative to the placebo treated control group. The dependent variables are available student-level individual characteristics and pre-treatment math scores (in the previous year endline tests). Gender is a dummy variable that switches on for a female student. Grades 1 to 6 are dummy variables that switch on for a student from grades one to six, respectively. Pre-Treatment Math Scores are standardized math scores held a year before our treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B1.4: Balance over Teacher and Student Characteristics by Mixed versus Same Gender Study Group

Panel A: Teacher Characteristics									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
	<i>Movie "Bol" Watched</i>	<i>Married</i>	<i>Av. Teaching Hours</i>	<i>Av. Class Size</i>	<i>Teaching Experience</i>	<i>Years of Education</i>	<i>Pre-Treatment Gender Index</i>	<i>Educational Specialization</i>	
<i>Mixed Study Group</i>	-0.0310 [0.0425]	0.0098 [0.0426]	0.0610 [0.2400]	-0.2920 [1.674]	0.1910 [0.2490]	0.0074 [0.1300]	0.0116 [0.0485]	-0.0059 [0.0373]	
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	607	607	607	607	607	607	607	607	
R-squared	0.072	0.072	0.059	0.086	0.071	0.054	0.099	0.082	
Mean of dependent var	0.532	0.473	30.28	24.81	4.608	12.679	-0.071	0.259	
Panel B: Student Characteristics									
	<i>Gender</i>	<i>Grade 1</i>	<i>Grade 2</i>	<i>Grade 3</i>	<i>Grade 4</i>	<i>Grade 5</i>	<i>Pre-Treatment Gender Index</i>	<i>Pre-Treatment Math Score</i>	<i>General Knowledge</i>
<i>Mixed Study Group</i>	-0.0090 [0.0159]	0.0143 [0.0296]	0.0240 [0.0299]	-0.0202 [0.0279]	0.0119 [0.0265]	-0.0080 [0.0276]	-0.0007 [0.0175]	-0.0424 [0.0386]	-0.0150 [0.0451]
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,932	13,932	13,932	13,932	13,932	13,932	13,932	13,932	13,932
R-squared	0.359	0.035	0.023	0.015	0.037	0.023	0.005	0.097	0.116
Mean dep. var.	0.511	0.163	0.173	0.153	0.144	0.131	0.000	0.000	0.000

Note: Robust standard errors appear in brackets (clustered at teacher level). The treatments are compared relative to the placebo treated control group. In Panel A, the dependent variables are teacher-level individual characteristics. Teaching Experience is years of teaching experience. Years of Education is the number of years of formal studies. Educational Specialization is a dummy variable that switches on for a specialization in Pedagogy. Av. Teaching Hours is the average number of hours taught every week. Av. Class Size is the average number of students in the class. Married is a dummy variable that switches on if a teacher is married. In Panel B, the dependent variables are available student-level individual characteristics and pre-treatment math scores (in the previous year endline tests). Gender is a dummy variable that switches on for a female student. Grades 1 to 6 are dummy variables that switch on for a student from grades one to six, respectively. Pre-Treatment Math Scores are standardized math scores held a year before our treatment. *Mixed Study Group* is a dummy variable that switches on when the students are assigned the mixed-gender study group. The outcomes are in levels to show the raw changes. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group.*** p<0.01, ** p<0.05, * p<0.1.

Table B2: Impact of Treatments on Social Behavior Games Played against Same Gender

	<i>Redistribution</i>	<i>Competitiveness</i>	<i>Cooperation</i>	<i>Coordination</i>
	(1)	(2)	(3)	(4)
<u><i>Movie X Mixed Study Group</i></u>	0.00175	-0.0502	-0.0410	-0.0508
p-value	(0.9784)	(0.4775)	(0.2940)	(0.3638)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.9980}	{0.9530}	{0.8012}	{0.8761}
<u><i>Movie-Curriculum X Mixed Study Group</i></u>	-0.0157	0.126*	-0.0564	0.0995*
p-value	(0.8162)	(0.0580)*	(0.1352)	(0.0854)*
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.9980}	{0.1838}	{0.4845}	{0.2877}
<u><i>U X Mixed Study Group</i></u>	-0.0781	-0.0171	0.0134	-0.0253
p-value	(0.2458)	(0.7954)	(0.7304)	(0.6587)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.7532}	{0.9980}	{0.9980}	{0.9930}
<u><i>M X Mixed Study Group</i></u>	-0.0354	-0.0844	0.0443	0.0141
p-value	(0.6042)	(0.2081)	(0.2580)	(0.7944)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.9860}	{0.6813}	{0.7602}	{0.9980}
Playing with Opposite Gender	Yes	Yes	Yes	Yes
Individual Controls & School FE	Yes	Yes	Yes	Yes
Observations	9,145	9,145	9,145	9,145
R-squared	0.010	0.009	0.640	0.012

Note: Dependent variables are outcomes on redistribution, competitiveness, cooperation, and coordination games, respectively standardized to mean zero and standard deviation one. The outcomes are recorded 12 months after the treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces (considering 16 hypotheses with 4 treatments X 4 outcomes). Further details on this are provided in Appendix D8 Each student plays the game twice, with the opposite gender (Table 5) and with the same gender (Table B2). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Table B3: Impact on Students' Stress

	<i>Stress Likert - Standardized</i>	
<i>Visual Narrative (movie)</i>	0.023 [0.034]	0.022 [0.035]
<i>Joint Movie-Curriculum</i>	0.024 [0.036]	0.025 [0.037]
School Fixed Effects	Yes	Yes
Individual Controls	No	Yes
Observations	8036	8036
R-squared	0.002	0.001
P-value (BM = BMC)	0.973	0.940

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. Dependent variable is standardized to mean zero and standard deviation one. The dependent variable is self-reported stress on a Likert Scale of 1-5 standardized to mean zero and standard deviation one. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The Utilitarian and Malleability treatments are also controlled for in both specifications. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores.*** p<0.01, **p<0.05, * p<0.1.

Table B4: Student Gender Attitudes by Grade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	<i>Kindergarten</i>	<i>Grade 1</i>	<i>Grade 2</i>	<i>Grade 3</i>	<i>Grade 4</i>	<i>Grade 5</i>	<i>Grade 6</i>
<i>Joint Movie-Curriculum</i>	0.248*	0.077	0.553***	0.501***	0.132	0.331**	0.174
	(0.143)	(0.143)	(0.121)	(0.137)	(0.140)	(0.132)	(0.154)
<i>Visual Narrative (movie)</i>	0.054	-0.039	0.359***	0.283**	-0.349**	0.011	0.360**
	(0.134)	(0.139)	(0.099)	(0.133)	(0.156)	(0.194)	(0.149)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3890	3747	4018	3528	3333	3022	1518
R-squared	0.146	0.101	0.110	0.120	0.090	0.155	0.117
P-value (BM = BMC)	0.143	0.463	0.082*	0.134	p < 0.01***	0.032**	0.352

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. The dependent variable is students' gender attitudes collected from Student Attitudinal Survey used in the paper. The outcomes are recorded 12 months after the treatment. The corresponding survey statements from students are reported in Appendix D4 in Appendix D. Visual narrative (movie) represents the standalone visual narrative treatment. Joint Movie and Curriculum is the dummy that switches on when the joint movie-curriculum treatment was turned on for the teacher. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student level controls include the dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five, six) and pre-treatment Math scores. School fixed effects are also included though the results are essentially identical without them. *** p<0.01, ** p<0.05, * p<0.1

Table B5: Heterogeneity by Student Gender

Panel A: Boys	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Gender Attitudes</i>	<i>Math Test Scores</i>	<i>Redistributio n</i>	<i>Competitivene ss</i>	<i>Cooperatio n</i>	<i>Coordination</i>
<i>Joint Movie-Curriculum</i>	-0.033 [0.035]	0.181*** [0.067]	-0.015 [0.032]	-0.011 [0.042]	-0.071 [0.124]	0.156*** [0.038]
<i>Visual narrative (movie)</i>	0.059 [0.037]	-0.097 [0.073]	0.022 [0.031]	0.045 [0.046]	-0.079 [0.125]	0.109*** [0.033]
<i>Utilitarian</i>	0.025 [0.037]	0.071 [0.080]	-0.030 [0.030]	0.010 [0.046]	0.086 [0.128]	0.006 [0.034]
<i>Malleability</i>	0.017 [0.035]	0.006 [0.076]	0.007 [0.032]	0.106** [0.048]	-0.024 [0.125]	-0.008 [0.030]
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11715	11705	9196	9196	9196	9196
R-squared	0.004	0.096	-0.001	0.008	0.072	0.005
Panel B: Girls	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Gender Attitudes</i>	<i>Math Test Scores</i>	<i>Redistributio n</i>	<i>Competitivene ss</i>	<i>Cooperatio n</i>	<i>Coordination</i>
<i>Joint Movie-Curriculum</i>	-0.049 [0.043]	0.150* [0.082]	0.007 [0.037]	-0.078 [0.053]	0.038 [0.136]	0.077* [0.041]
<i>Visual narrative (movie)</i>	-0.060 [0.039]	0.085 [0.082]	-0.043 [0.036]	-0.006 [0.045]	-0.098 [0.131]	0.020 [0.034]
<i>Utilitarian</i>	-0.073* [0.040]	0.157* [0.083]	-0.044 [0.030]	-0.015 [0.046]	-0.167 [0.128]	0.060** [0.030]
<i>Malleability</i>	0.001 [0.041]	0.019 [0.085]	-0.012 [0.035]	0.006 [0.048]	-0.004 [0.139]	-0.001 [0.034]
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11361	11350	9094	9094	9094	9094
R-squared	0.006	0.095	-0.002	0.006	0.039	-0.001

Note: Dependent variables are outcomes on student gender attitudes, math test scores, behavioral games such as redistribution, competitiveness, cooperation, and coordination, standardized to mean zero and standard deviation one. The outcomes are recorded 12 months after the treatment. Panel A presents the analysis on boys, while Panel B presents the analysis on girls. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. Each student plays the game twice, with the same gender (Table 5) and with the opposite gender (Table B2). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. Standard errors are clustered at the teacher level. *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Robustness - Marlowe-Crowne Social Desirability

	(1)	(2)	(3)	(4)
	<i>Women's Rights Overall</i>	<i>Petition to Criminalize Dowry</i>	<i>Petition to Abolish Polygamy</i>	<i>IAT Score</i>
<i>Visual narrative (movie)</i>	0.171*** [0.0577]	0.0947** [0.0464]	0.0749*** [0.0288]	0.245* [0.136]
<i>Joint Movie-Curriculum</i>	0.225*** [0.0587]	0.157*** [0.0497]	0.0930*** [0.0281]	0.346** [0.162]
<i>Utilitarian</i>	0.0737 [0.0531]	-0.0113 [0.0383]	0.00174 [0.0116]	-0.0789 [0.140]
<i>Malleability</i>	0.121** [0.0593]	0.0005 [0.0396]	-0.00491 [0.131]	-0.123 [0.123]
Individual controls	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	526	526	526	526
R-squared	0.162	0.144	0.21	0.131
P-value (BM = BMC)	0.356	0.26	0.633	0.538

Note: The dependent variables are identical to those used in the regressions in the main text. The treatments are compared relative to the placebo treated control group. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. Specifically, we discard individuals who always answer yes to the following statements: 1) I am never jealous of another person's good fortune 2) I am always a good listener. 3) I am never angry. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1

Table B7: Impact on Gender Attitudes and Student Achievement - Assessing Spillovers

Panel A: Teacher Characteristics	(1)	(2)	(3)	(4)
	<i>Women's Rights Overall</i>	<i>Petition to Criminalize Dowry</i>	<i>Petition to Abolish Polygamy</i>	<i>Standardized Gender IAT Score</i>
<i>Fraction of Joint Treated</i>	-0.204	0.167	-0.0329	-1.355
<i>Teachers X Joint Treatment</i>	[0.332]	[0.321]	[0.169]	[1.232]
<i>Joint Movie-Curriculum</i>	0.242**	0.135	0.0957*	0.693*
	[0.106]	[0.0955]	[0.0564]	[0.374]
Individual controls	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	607	607	607	527
R-squared	0.139	0.141	0.201	0.134
Panel B: Student Characteristics	(1)	(2)	(3)	(4)
	<i>Student Attitudinal Survey</i>		<i>Mathematics Scores</i>	
<i>Fraction of Joint Treated</i>	-0.138	-0.121	0.180	0.158
<i>Teachers X Joint Treatment</i>	[0.308]	[0.306]	[0.327]	[0.165]
<i>Joint Movie-Curriculum</i>	0.290***	0.277***	0.113	0.0772
	[0.0888]	[0.0885]	[0.106]	[0.0544]
Individual controls	No	Yes	No	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	13,911	13,911	13,911	13,911
R-squared	0.038	0.044	0.090	0.596

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. The dependent variable in Column (1) is an index consisting of 16 gender rights statements fielded concerning Women's Economic, Social, Legal and Political Rights. The statements can be found in Appendix D2. In Column 2, the dependent variable is a petition sent to the Pakistani parliament to criminalize dowry, while the dependent variable in Column (3) is a petition to abolish the law that allows polygamy in Pakistan. Column (4) has the dependent variable on gender Implicit Association Test (IAT). All dependent variables are standardized to mean zero and standard deviation one. The Fraction of Joint Treated Teachers is the proportion of teachers treated with the joint treatment within schools. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student level controls include the dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five, six) and pre-treatment Math scores. School fixed effects are also included. *** p<0.01, ** p<0.05, * p<0.1

**Table B8: Impact of schools more intensely treated by the Visual Narrative
on Gender Attitudes**

Panel A: Utilitarian Treated Group						
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Gender Rights Index</i>		<i>Petition</i>		<i>Gender IAT Score</i>	
<i>Fraction Treated with the Visual Narrative</i>	0.00558 [0.0107]	0.00366 [0.0103]	0.00447 [0.0124]	0.00498 [0.0115]	-0.0686 [0.0476]	-0.0736 [0.0575]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	121	121	121	121	107	107
R-squared	0.002	0.041	0.002	0.082	0.022	0.101
Panel B: Malleability Treated Group						
<i>Fraction Treated with the Visual Narrative</i>	0.00753 [0.0153]	0.00967 [0.0151]	0.000161 [0.0066]	0.00144 [0.0058]	-0.00493 [0.0258]	-0.00749 [0.0259]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	121	121	121	121	121	121
R-squared	0.002	0.042	0.000	0.031	0.000	0.035
Panel C: Placebo Treated Group						
<i>Fraction Treated with the Visual Narrative</i>	-0.0153 [0.0126]	-0.0157 [0.0121]	-0.00409 [0.0075]	-0.00329 [0.0085]	-0.0234 [0.0315]	-0.0198 [0.0327]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	122	122	122	122	85	85
R-squared	0.010	0.104	0.002	0.063	0.003	0.094

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. The dependent variable in Column 1 and 2 is an index consisting of 16 gender rights statements fielded concerning Women’s Economic, Social, Legal and Political Rights. The statements can be found in Appendix D2. In Column 3 and 4, the dependent variable is a signed petition sent to the Pakistani parliament to criminalize dowry, Column 5 and 6 estimates the main specification with the gender Implicit Association Test (IAT) as the dependent variable. All dependent variables in this table are *standardized* to mean zero and standard deviation one. Panel A represents the Utilitarian treated group, Panel B represents the Malleability treated group and Panel C represents the Placebo treated group. The *Fraction of Visual Narrative Treated Teachers* is the proportion of teachers treated with the *Visual Narrative* treatment within schools. *Visual narrative* represents the visual narrative treatment of the movie Bol with a structured discussion of gender rights themes in the movie. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student level controls include the dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five, six) and students’ pre-treatment gender rights index. *** p<0.01, ** p<0.05, * p<0.1

Table B9: Impact of schools more intensely treated by the Joint Treatment on Gender Attitudes

Panel A: Utilitarian Treated Group						
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Gender Rights Index</i>		<i>Petition</i>		<i>Gender IAT Score</i>	
<i>Fraction Treated with</i>	0.489*	0.532*	-0.112	-0.122	-0.340	-0.130
<i>Joint Movie-Curriculum</i>	[0.264]	[0.274]	[0.185]	[0.189]	[0.785]	[0.757]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	121	121	121	121	107	107
R-squared	0.033	0.078	0.003	0.084	0.001	0.076
Panel B: Malleability Treated Group						
<i>Fraction Treated with</i>	-0.0490	-0.0542	-0.0851	-0.0626	0.930	0.932
<i>Joint Movie-Curriculum</i>	[0.222]	[0.243]	[0.147]	[0.150]	[0.624]	[0.699]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	121	121	121	121	121	121
R-squared	0.000	0.038	0.002	0.032	0.028	0.060
Panel C: Placebo Treated Group						
<i>Fraction Treated with</i>	-0.170	-0.233	0.153	0.213	-0.0417	0.383
<i>Joint Movie-Curriculum</i>	[0.253]	[0.274]	[0.180]	[0.170]	[0.723]	[0.746]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	122	122	122	122	85	85
R-squared	0.004	0.102	0.007	0.075	0.000	0.095

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. The dependent variable in Column 1 and 2 is an index consisting of 16 gender rights statements fielded concerning Women’s Economic, Social, Legal and Political Rights. The statements can be found in Appendix D2. In Column 3 and 4, the dependent variable is a signed petition sent to the Pakistani parliament to criminalize dowry, Column 5 and 6 estimates the main specification with the gender Implicit Association Test (IAT) as the dependent variable. All dependent variables in this table are standardized to mean zero and standard deviation one. Panel A presents the analysis on the Utilitarian treated teachers, Panel B the Malleability treated teachers and Panel C the Placebo treated group. The *Fraction of Joint Treated Teachers* is the proportion of teachers treated with the joint treatment within schools. *Joint* represents the visual narrative treatment of the movie Bol with a structured discussion of gender rights themes in the movie. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student level controls include the dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five, six) and students’ pre-treatment gender rights index. *** p<0.01, ** p<0.05, * p<0.1.

Table B10: Impact of Previously Bol Fraction of Teachers on Gender Attitudes in Control Group

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Gender Rights Index</i>		<i>Petition</i>		<i>Gender IAT Score</i>	
<i>Fraction of Teachers Who Previously Watched Bol Movie</i>	-0.143 [0.229]	-0.0786 [0.232]	0.155 [0.188]	0.159 [0.194]	0.044 [0.574]	0.273 [0.586]
Individual Controls	No	Yes	No	Yes	No	Yes
Observations	122	122	122	122	85	85
R-squared	0.003	0.065	0.007	0.065	0.000	0.091

Note: Robust standard errors appear in brackets (clustered at the teacher level). The treatments are compared relative to the placebo treated control group. The dependent variable in Column 1 and 2 is an index consisting of 16 gender rights statements fielded concerning Women’s Economic, Social, Legal and Political Rights. The statements can be found in Appendix D2. In Column 3 and 4, the dependent variable is a dummy variable that switches on if the petition is sent to the Pakistani parliament to criminalize dowry, Column 5 and 6 estimates the main specification with the gender Implicit Association Test (IAT) as the dependent variable. All dependent variables in this table are standardized to mean zero and standard deviation one. Table presents the analysis on the Placebo treated group. The Fraction of Previously Watched Bol Movie is the proportion of teachers in the control group who previously watched ‘Bol’ movie. The teacher level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student level controls include the dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five, six) and students’ pre-treatment gender rights index. *** p<0.01, ** p<0.05, * p<0.1.

Table B11: Impact of Treatment on Social Behavior - Same Gender Study

Group

	<i>Redistribution</i>	<i>Competitiveness</i>	<i>Cooperation</i>	<i>Coordination</i>
	(1)	(2)	(3)	(4)
<u>Visual Narrative (Movie)</u>	-0.0287	0.0590	-0.0124	0.0109
p-value	(0.4031)	(0.2189)	(0.5843)	(0.5994)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.961}	{0.8361}	{0.9920}	{0.9920}
<u>Joint Movie-Curriculum</u>	0.00190	-0.0546	-0.0214	0.0185
p-value	(0.9569)	(0.2368)	(0.3295)	(0.3511)
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.999}	{0.860}	{0.945}	{0.9460}
<u>Utilitarian</u>	-0.00327	0.0320	-0.0188	0.0435
p-value	(0.9199)	(0.4509)	(0.3956)	(0.0321)**
Sharpened q-value	[0.999]	[0.999]	[0.999]	[0.318]
Romano-Wolf corrected p-value	{0.999}	{0.967}	{0.961}	{0.176}
<u>Malleability</u>	-0.00475	0.130	0.00262	-0.000158
p-value	(0.886)	(0.0047)***	(0.914)	(0.9937)
Sharpened q-value	[0.999]	[0.082]*	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.999}	{0.0240}**	{0.999}	{0.999}
Playing with Opposite Gender	Yes	Yes	Yes	Yes
Individual Controls & School FE	Yes	Yes	Yes	Yes
Observations	9,048	9,048	9,048	9,048
R-squared	0.007	0.020	0.010	0.003

Note: Dependent variables are outcomes on redistribution, competitiveness, cooperation, and coordination games, respectively standardized to mean zero and standard deviation one. The outcomes are recorded 12 months after the treatment. The outcomes are recorded 12 months after the treatment. 'Utilitarian' variable is a binary indicator that assumes a value of one upon the teacher's receipt of the corresponding treatment, analogous to the 'Malleability' treatment indicator. 'Visual Narrative (Movie)' is similarly a dummy turning on for subjects assigned the Bol Movie. The 'Joint Movie and Curriculum' indicator turns on for teachers assigned the joint Bol Movie and the Gender-rights Curriculum treatment. Each treatment is followed by a 30-minute structured discussion, the particulars of which are delineated in Table A3. The treatments are compared relative to the placebo treated control group. The treatments are compared relative to the placebo treated control group. Each student is playing against the Opposite Gender in this table, as also emphasized by a row in the table. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. P-values computed using the Newey-West estimator are reported in parentheses, along with the multiple hypothesis-adjusted FDR q-values in square brackets and FWER-adjusted p-values in curly braces (considering 16 hypotheses with 4 treatments X 4 outcomes). The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Appendix C. Short-Run Results - 6 Months Post-Treatment

Table C1: Impact on Teachers' Attitude - Short-Term Results

<i>Women's:</i>	<i>Rights Overall</i>	<i>Economic Rights</i>	<i>Political Rights</i>	<i>Social Rights</i>	<i>Legal Rights</i>
	(1)	(2)	(3)	(4)	(5)
<u><i>Visual narrative (movie)</i></u>	0.129	0.152	0.222	0.0252	0.210
p-value	(0.0365)**	(0.0267)**	(0.0774)*	(0.7323)	(0.0875)*
Sharpened q-value	[0.108]	[0.135]	[0.157]	[0.628]	[0.158]
Romano-Wolf corrected p-value	{0.061}*	{0.048}**	{0.1728}	{0.920}	{0.1958}
<u><i>Joint Movie-Curriculum</i></u>	0.208	0.241	0.389	0.0474	0.312
p-value	(0.0007)***	(0.001)***	(0.002)***	(0.5238)	(0.0086)***
Sharpened q-value	[0.011]**	[0.094]*	[0.013]**	[0.487]	[0.038]**
Romano-Wolf corrected p-value	{0.001}***	{0.001}***	{0.002}***	{0.8651}	{0.012}**
<u><i>Utilitarian</i></u>	0.0542	0.0579	0.0571	0.0780	-0.0209
p-value	(0.2969)	(0.3654)	(0.5702)	(0.2558)	(0.8236)
Sharpened q-value	[0.296]	[0.354]	[0.505]	[0.291]	[0.701]
Romano-Wolf corrected p-value	{0.6763}	{0.7722}	{0.8691}	{0.6543}	{0.920}
<u><i>Malleability</i></u>	0.0984	0.132	0.153	0.0306	0.0959
p-value	(0.1125)	(0.0553)*	(0.2544)	(0.6688)	(0.4464)
Sharpened q-value	[0.187]	[0.135]	[0.291]	[0.592]	[0.424]
Romano-Wolf corrected p-value	{0.2537}	{0.1179}	{0.6543}	{0.9201}	{0.8531}
Individual controls	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	607	607	607	607	607
R-squared	0.134	0.111	0.126	0.116	0.137
P-value BM = BMC	.226	.251	.231	.767	.455

Note: The figure summarizes our main results – effect of the treatments on the different indices summarizing attitudes to women rights. The treatments are compared relative to the placebo treated control group. The outcomes are recorded 6 months after the treatment. Women's Rights Overall is an index consisting of all the statements concerning Women's Economic, Social, Legal and Political Rights. Women's Economic Rights is an index combining women's rights to education and work outside home, based on reactions to statements "Women should be allowed to work outside the home". "Women and men should have equal rights to jobs". "I have no problem with my sister or female cousin working outside the home". "Daughters should have the same right to inherit property as sons". "Women and men should have equal rights to get an education". "Wives should not be less educated than their husbands". "Boys should not have more opportunities and resources for education than girls." Women's Political Rights is based on statements "It would be a good idea to elect a woman as the village Sarpanch (local politician)." "Women and men have equal rights to be President or Prime Minister." Women's Social Rights is based on statements "Domestic violence by husbands cannot be justified" "Parents should seek their daughter's consent before fixing her marriage". "A woman should not necessarily get married before her 25th Birthday". "Women who give birth to a son need not be honored in the family". "A woman with five daughters should not be under social pressure to bear a son." Finally, the Women's Legal Rights index is based on statements "Laws should be passed to ban dowry." "Under Article 35 of the Constitution of Pakistan & Judgment of Federal Shariat Court, the consent of 'Wali' is not required and a sui juris Muslim female can enter into a valid Nikah / Marriage under her own free will without the consent of Wali. To what extent do you approve of this legal right of women to enter marriage under their own free will." Equation (1) is estimated with all controls and explanatory variables. Utilitarian is a dummy variable that switches on when the teacher receives this type of treatment, same for Malleability treatment. Visual narrative is a dummy that switches on for participants receiving the Bol Movie treatment. Joint Movie and Curriculum is a dummy that switches on for participants receiving the Bol Movie and Gender-rights Curriculum treatment. P-values are robust and presented with the MHT FDR and FWER corrections described in Appendix D8 (considering 20 hypotheses with 4 treatments X 5 outcomes). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores.*** p<0.01, ** p<0.05, * p<0.1.

Table C2: Impact on Standardized Student Math Test Scores by Students Gender - Short-Term Results

	<i>Math Test Scores</i>				
	(1)	(2)	(3)	(4)	(5)
<i>Visual narrative (movie)</i>	-0.0426	-0.00777	-0.0205	0.00319	-0.0013
p-value	(0.2626)	(0.8116)	(0.518)	(0.9174)	(0.9657)
Sharpened q-value	[0.829]	[0.999]	[0.999]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.3826}	{0.988}	{0.7842}	{0.999}	{0.999}
<i>Joint Movie-Curriculum</i>	0.175***	0.0467	0.159***	0.104***	0.0584*
p-value	p < 0.01***	(0.1993)	p < 0.01***	(0.001)***	(0.0754)*
Sharpened q-value	[0.001]***	[0.829]	[0.001]***	[0.008]***	[0.472]
Romano-Wolf corrected p-value	{0.0010}***	{0.2168}	{0.001}***	{0.001}***	{0.026}**
<i>Utilitarian</i>	0.0384	0.0457	0.0322	0.0323	0.0449
p-value	(0.2792)	(0.1854)	(0.3398)	(0.3337)	(0.1835)
Sharpened q-value	[0.829]	[0.829]	[0.829]	[0.829]	[0.829]
Romano-Wolf corrected p-value	{0.4026}	{0.2088}	{0.4845}	{0.4845}	{0.2088}
<i>Malleability</i>	-0.000233	-0.0268	0.0300	-0.0136	-0.001
p-value	(0.9946)	(0.4411)	(0.318)	(0.6632)	(0.9853)
Sharpened q-value	[0.999]	[0.999]	[0.829]	[0.999]	[0.999]
Romano-Wolf corrected p-value	{0.999}	{0.6733}	{0.4785}	{0.9191}	{0.999}
<i>Movie-Curriculum X Mixed Study</i>					0.139*** (0.0461)
<i>Movie X Mixed Study Group</i>					-0.0255 (0.0480)
<i>U X Mixed Study Group</i>					-0.0239 (0.0482)
<i>M X Mixed Study Group</i>					0.005 (0.0462)
<i>Mixed Study Group</i>					0.0105 (0.0302)
Students are Girls			No	Yes	
Mixed-gender Study Group Sample	Yes	No			
Individual Controls & School FE	Yes	Yes	Yes	Yes	Yes
Observations	6,973	6,959	7,117	6,815	13,932
P-value BM = BMC	<0.001***	0.132	<0.001***	0.002***	0.082*

Note: The dependent variable is the standardized math scores. The treatments are compared relative to the placebo treated control group. The outcomes are recorded 6 months after the treatment. Utilitarian is a dummy variable that switches on when the teacher receives this type of treatment, same for Malleability treatment. Visual narrative is a dummy that switches on for participants receiving the Bol Movie treatment. Joint Movie and Curriculum is a dummy that switches on for participants receiving the Bol Movie and Gender-rights Curriculum treatment. P-values are robust and presented with the MHT FDR and FWER corrections described in Appendix D8 (considering 20 hypotheses with 4 treatments X 5 outcomes). Mixed Study Group is a dummy that switches on when the student study group is mixed-gender. *U X Mixed Study Group*, *M X Mixed Study Group*, *Movie X Mixed Study Group*, and *Joint Movie-Curriculum X Mixed Study Group* are interaction terms of Mixed Study Group with U, M, movie, and joint movie-curriculum treatments, respectively. The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include the dummies for student grade and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Impact of Teachers’ Training on Students’ Attitudes and Math Test Scores - Short-Term Results

	<i>Student Attitudinal Survey</i>		<i>Maths</i>	
	(1)	(2)	(3)	(4)
<u>Visual narrative (movie)</u>	0.107*	0.105**	-0.0136	-0.0124
p-value	(0.0754)*	(0.0479)**	(0.8374)	(0.6001)
Sharpened q-value	[0.178]	[0.13]	[0.772]	[0.392]
Romano-Wolf corrected p-value	{0.005}***	{0.002}***	{0.963}	{0.6464}
<u>Joint Movie-Curriculum</u>	0.276***	0.278***	0.178***	0.129***
p-value	p < 0.01***	p < 0.01***	(0.0049)***	p < 0.01***
Sharpened q-value	[0.001]***	[0.001]***	[0.017]***	[0.001]***
Romano-Wolf corrected p-value	{0.001}***	{0.001}***	{0.001}***	{0.001}***
<u>Utilitarian</u>	-0.0554	-0.0944	0.113	0.0320
p-value	(0.4324)	(0.1111)	(0.1001)	(0.1987)
Sharpened q-value	[0.459]	[0.182]	[0.182]	[0.284]
Romano-Wolf corrected p-value	{0.2787}	{0.011}**	{0.01}**	{0.049}**
<u>Malleability</u>	-0.00359	0.00128	-0.000381	0.00102
p-value	(0.9585)	(0.9831)	(0.9954)	(0.9657)
Sharpened q-value	[0.772]	[0.772]	[0.772]	[0.772]
Romano-Wolf corrected p-value	{0.997}	{0.999}	{0.999}	{0.999}
Individual controls	No	Yes	No	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Observations	13,932	13,932	13,932	13,932
R-squared	0.084	0.183	0.108	0.716
P-value BM = BMC	<0.001***	<0.001***	0.0045***	<0.001***

Note: The dependent variables are standardized to mean zero and standard deviation for Math scores and attitudinal. The treatments are compared relative to the placebo treated control group. The outcomes are recorded 6 months after the treatment. The corresponding survey statements from students are reported in Appendix D4. Utilitarian is a dummy variable that switches on when the teacher receives this type of treatment, same for Malleability treatment. Visual narrative is a dummy that switches on for participants receiving the Bol Movie treatment. Joint Movie and Curriculum is a dummy that switches on for participants receiving the Bol Movie and Gender-rights Curriculum treatment. P-values are robust and presented with the MHT FDR and FWER corrections described in Appendix D8 (considering 16 hypotheses with 4 treatments X 4 outcomes). The teacher-level controls include years of teaching experience, educational qualification, professional qualification, average teaching hours, class size, and marital status. The student-level controls include dummies for student grade (i.e., KG, Nursery, Prep, one, two, three, four, five and six class) and pre-treatment math scores. *** p<0.01, ** p<0.05, * p<0.1.

D. Data Appendix: Survey and Data Details

Appendix D1. Consent

For teachers:

I agreed to participate in the research study. I understand the purpose and nature of this study and I am participating voluntarily. I understand that I can withdraw from the study at any time, without any penalty or consequences.

Yes No

I grant permission for the data generated from this survey to be used in the researcher's publications on this topic.

Yes No

I grant permission to researchers to use my anonymized information for research purposes and this includes my personal data with PEN.

Yes No

For parents:

I grant permission to researchers to use my son or daughter's anonymized information for research purposes and this includes the personal data with PEN.

Yes No

Appendix D2. Gender Rights Survey fielded to Teachers

Likert Scale:

1. Totally Disagree
2. Disagree
3. Neutral
4. Agree
5. Totally Agree

S1. Women should be allowed to work outside the home.

S2. Women and men should have equal rights to jobs.

S3. I have no problem with my sister or female cousin from working outside the home.

S4. Daughters should have a similar right to inherited property as sons.

S5. Women and men should have equal rights to get an education as men.

S6. Wives should not be less educated than their husbands.

S7. Boys should not get more opportunities and resources for education than girls

S8. It would be a good idea to elect a woman as the village Sarpanch (local politician).

S9. Women and men have equal rights to be President or Prime Minister.

S10. Domestic violence by husbands cannot be justified.

S12. Women should not necessarily get married before her 25th Birthday.

S13. Women who give birth to a son need not be honored in the family.

S14. A woman with five daughters should not be under social pressure to bear a son.

S15. Laws should be passed to ban dowry.

S16. Under Article 35 of the Constitution of Pakistan & Judgment of Federal Shariat Court, the consent of 'Wali' is not required and a sui juris Muslim female can enter into a valid Nikah / Marriage under her own freewill without the consent of Wali. How much do you approve of this legal right of women to enter marriage under their own freewill.

Appendix D3. Procedure for Index Construction

Gender rights index averages across all components of gender rights survey questions listed previously. Women's Rights Overall is an index consisting of all the statements concerning Women's Economic, Social, Legal and Political Rights i.e. all the 16 statements in section D2. Women's Economic Rights is an index combining women rights relevant to education and work outside home i.e. statements 1 to 7. Women's Political rights is an index of statements 8 and 9, while women's social rights is based on statements 10 to 14. Finally, the legal rights index combines statements 15 and 16.

Specifically, Women's Rights Overall is an index consisting of all the statements concerning Women's Economic, Social, Legal and Political Rights. Women's Economic Rights is an index combining women rights relevant to education and work outside home i.e. statements "Women should be allowed to work outside the home". "Women and men should have equal rights to jobs". "I have no problem with my sister or female cousin from working outside the

home”. “Daughters should have a similar right to inherited property as sons”. “Women and men should have equal rights to get an education as men”. “Wives should not be less educated than their husbands”. “Boys should not get more opportunities and resources for education than girls.”. Women’s Political rights is an index of statements “It would be a good idea to elect a woman as the village Sarpanch (local politician).” “Women and men have equal rights to be President or Prime Minister.”, while women's social rights index is based on statements “Domestic violence by husbands cannot be justified” “Parents should seek their daughter's consent before fixing her marriage”. “Women should not necessarily get married before her 25th Birthday”. “Women who give birth to a son need not be honored in the family”. “A woman with five daughters should not be under social pressure to bear a son.”. Finally, the legal rights index combines statements “Laws should be passed to ban dowry. Under Article 35 of the Constitution of Pakistan & Judgment of Federal Shariat Court, the consent of ‘Wali’ is not required and a sui juris Muslim female can enter into a valid Nikah / Marriage under her own freewill without the consent of Wali. How much do you approve of this legal right of women to enter marriage under their own freewill.”

Appendix D4. Petition Details and Template Presented to all teachers

The petitions are likely high-stakes for two reasons. First, sending a petition to parliament with teachers’ names and national identity card numbers in a conservative society amounts to a revealed preference measure of gender attitudes and can be seen as women exerting their political rights. Second, such a gesture may have economic, social, and political repercussions for public school teachers. We administer this via attaching two blank petitions with our endline survey that gives an opportunity to sign a petition to the Pakistani parliament to both treated and control teachers to abolish polygamy and criminalize dowry. We then send all the petitions to the parliamentary committee on protection of women

at the National Assembly of Pakistan where debates on these bills are currently taking place. The empty petition text was provided to all teachers in a separate room with our instructions attempting to reduce experimental demand by explicitly mentioning that the teachers are free not to sign the petition and leave it blank if they wish to do so. The text of the blank petitions with exact instructions that were provided to all teachers, including those in the control group, can be found in this appendix. The specific petitions provided with instructions are below.

Petition Instructions and Text

Please fill this in a separate room individually. Please also note that we will actually send this petition to the National Assembly of Pakistan, so feel free to leave one or both petitions blank if you wish not to send one or both of these petitions.

Polygamy Petition

I, ----- (enter full name), daughter of -----, am signing this petition to request the complete ban on polygamy.

I hereby grant permission to send this petition to demand from the National Assembly and Senate of Pakistan to repeal the Muslim family law pertaining to polygamy. The law is as follows:

“6. Polygamy.– (1) No man, during the subsistence of an existing marriage, shall, except with the previous permission in writing of the Arbitration Council, contract another marriage, nor shall any such marriage contracted without such permission be registered under this Ordinance.

(2) An application for permission under sub-section (1) shall be submitted to the Chairman in the prescribed manner, together with the prescribed fee, and shall state reasons for the proposed marriage, and whether the consent of existing wife or wives has been obtained thereto.

(3) On receipt of the application under sub-section (2), the Chairman shall ask the applicant and his existing wife or wives each to nominate a representative, and the Arbitration Council so constituted may, if satisfied that the proposed marriage is necessary and just, grant, subject to such conditions, if any, as may be deemed fit, the permission applied for.”

Yours Truly,

Your CNIC:

Your Full Name:

Dowry Petition

I, ----- (enter full name), daughter of -----, am signing this petition to request the complete ban and criminalization of dowry.

I hereby grant permission to send this petition to demand from the National Assembly and Senate of Pakistan to make dowry a criminal offense pertaining imprisonment for up to three years.

Yours Truly,

CNIC:

Your Full Name:

Appendix D5. IAT Test Details. Gender IAT in Urdu with an English translation

Press "E"

Woman

Press "I"

Men

1

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for the kind words **she**.

Press the key with your right index finger. Words for "**I**" Type **Male**.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

1

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے:

کے لئے کے الفاظ قسم **عورت** "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر
کے لئے کے الفاظ قسم **مرد** "I" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

ایک غلطی کے دوران، صحیح کی کو دبائے سے غلطی کو درست کریں۔ **X** ظاہر ہوتا ہے
جتنی جلدی ہو سکے جائز ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے **space bar** پر کلک کریں

Press "E"

Family

Press "I"

Career

2

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for the kind words **Family**.

Press the key with your right index finger. Words for "**I**" Type **Career**.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

2

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے:

کے لئے کے الفاظ قسم **کنبہ "E"** دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر
کے لئے کے الفاظ قسم **کیریئر "I"** اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

ایک غلطی کے دوران، صحیح کی کو دبائے سے غلطی کو درست کریں۔ **X** ظاہر ہوتا ہے
جتنی جلدی ہو سکے جائز ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے **space bar** پر کلک کریں

Press "E"

Woman

○

Family

Press "I"

Men

○

Career

3

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for Woman y Family.

Press the key with your right index finger. for "**I**" Men y Career.

Each word belongs to only one type.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar to** start.

دبائیں "I" پر

مرد

O

کیریئر

دبائیں "E" پر

عورت

O

کنہ

3

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے

کنہ y کے لئے عورت "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر

کیریئر y کے لئے مرد "I" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

ہر لفظ صرف ایک ہی سے تعلق رکھتا ہے قسم

ایک غلطی کے دوران، صحیح کی کو دبائے سے غلطی کو درست کریں۔ X ظاہر ہوتا ہے

جتنی جلدی ہو سکے جائز ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے space bar پر کلک کریں

Press "E"

Woman

○

Family

Press "I"

Men

○

Career

4

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for **Woman** y **Family**.

Press the key with your right index finger. for "**I**" **Men** y **Career**.

Each word belongs to only one type.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

دبائیں "I" پر

مرد

O

کیریئر

دبائیں "E" پر

عورت

O

کنہ

4

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے

کنہ y کے لئے عورت "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر

کیریئر y کے لئے مرد "I" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

ہر لفظ صرف ایک ہی سے تعلق رکھتا ہے قسم

ایک غلطی کے دوران، صحیح کی کو دبائے سے غلطی کو درست کریں۔ X ظاہر ہوتا ہے

جتنی جلدی ہو سکے جاؤ ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے space bar پر کلک کریں

Press "E"

Men

Press "I"

Woman

5

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for the kind words **Men**.

Press the key with your right index finger. Words for "**I**" Type **Woman**.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

دبائیں "ا" پر
عورت

دبائیں "E" پر
مرد

5

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے:

کے لئے کے الفاظ قسم مرد "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر
کے لئے کے الفاظ قسم عورت "ا" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

ایک غلطی کے دوران صحیح کی کو دبائے سے غلطی کو درست کریں۔ X ظاہر ہوتا ہے
جتنی جلدی ہو سکے جاؤ ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے space bar پر کلک کریں

Press "E"

Men

○

Family

Press "I"

Woman

○

Career

6

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for **Men** y **Family**.

Press the key with your right index finger. for "**I**" **Woman** y **Career**.

Each word belongs to only one type.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

دبائیں "E" پر

مرد

O

کنبہ

دبائیں "I" پر

عورت

O

کیریئر

6

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے

کنبہ y کے لئے مرد "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر

کیریئر y کے لئے عورت "I" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔

پر لفظ صرف ایک ہی سے تعلق رکھتا ہے قسم

ایک غلطی کے دوران۔ صحیح کی کو دبائے سے غلطی کو درست کریں۔ X ظاہر ہوتا ہے

جتنی جلدی ہو سکے جاؤ ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے space bar پر کلک کریں

Press "E"

Men

○

Family

Press "I"

Woman

○

Career

7

This page contains instructions. The game starts on the next page.

No word appears:

Press your finger left key "**E**" for **Men** y **Family**.

Press the key with your right index finger, for "**I**" **Woman** y **Career**.

Each word belongs to only one type.

X appears during an error. Correct the error by pressing the correct key.

Go as fast as possible while making some mistakes.

Click **space bar** to start.

دبائیں "E" پر مرد O کنبہ	دبائیں "I" پر عورت O کیریئر
-----------------------------------	--------------------------------------

Z

اس صفحے میں ہدایات ہیں۔ کھیل اگلے صفحے پر شروع ہوتا ہے۔

جب کوئی لفظ ظاہر ہوتا ہے:

کنبہ y کے لئے مرد "E" دبائیں آپ کی بائیں طرف کی انگلی کی چابی پر
کیریئر y کے لئے عورت "I" اپنی دائیں شہادت کی انگلی سے کلید دبائیں۔
ہر لفظ صرف ایک ہی سے تعلق رکھتا ہے قسم

ایک غلطی کے دوران، صحیح کی کو دبائے سے غلطی کو درست کریں۔ X ظاہر ہوتا ہے
جتنی جلدی ہو سکے جاؤ ممکن ہو سکے کے طور پر کچھ غلطیاں کرتے ہوئے

شروع کرنے کے لئے space bar پر کلک کریں

Appendix D6. General Knowledge Exam Embedded Survey on Gender Attitudes for Students

Likert Scale:

1. Totally Disagree
2. Disagree
3. Neutral
4. Agree
5. Totally Agree

S1. I think my mama should be allowed to work outside the home.

S2. Boys and girls should get the same opportunities to study.

S3. Girls should not be less educated than boys.

S4. Girls like boys can also become pilots.

S5. My mama can be the school principal “headmaster”.

Appendix D7. Details on Strategic Dilemmas with Students

Instructions. You will participate in four activities that will give you the opportunity to win milk carton(s). You will win milk carton coupons which you can exchange at your school canteen. Each credit you win will allow you to win a milk carton. If at any point, you feel confused or need help about instructions of the activity, feel free to reach out to one of the helping staff who will be happy to explain the activity. Our staff is only here to help you. We now request you to open the piece of papers provided to you and begin your activities. Good luck.

1. Cooperation Game

In this game, you play with another participant **for 4 turns**.

You keep the same partner during the 4 turns.

[4 variations:

Condition 1: The second participant is called [name]

Condition 2: The second participant is also in your class.

Condition 3: The second participant comes from another school.

Condition 4: The second participant comes from another class.

Note: the identity of the second player changes each turn.

Each turn, you both receive an initial endowment of 10 credit that can be changed into a milk carton at your canteen (1 credit = 1 milk coupon)

Decision on your part

You must decide how much of this initial endowment you wish to transfer to the other participant (between 0 and 10 credit). The transferred quantity will be doubled, and the other participant will receive this doubled quantity. What you choose not to transfer remains in your possession but however will not be doubled.

Decision (simultaneous) from your partner

The other participant simultaneously makes the same decision. He decides how much of his initial endowment he wishes to transfer to you (between 0 and 10 credit). You will receive double this transferred amount.

Your winnings on a round are calculated as the sum of what you keep (from your initial

endowment) plus double what the other participant transfers to you. The helping staff will help you with this.

At the end of each round, you will be able to know the decision made by the other participant and how many credits you have won on that round.

The First round will start on the next page.

Please choose how much of your initial endowment (between 0 and 10 credit) you wish to transfer to the other participant.

Please use a multiple of 1 credits: ____ credit (s)

You can view the score of round 1 on the Next page. The helping staff will fill it for you.

[Next Page: The helping staff will fill this out for you

You have transferred ___ credits to the other participant on this round.

The other participant transferred ___ credits to you.

This means that you have earned ___ credits and the other participant has earned ___ credits.

*This game will repeat for 3 more rounds

2. **Coordination Game**

In this game, you will interact with another participant for 4 rounds.

4 variations:

Condition 1: The second participant is called [name]...

Condition 2: The second participant is also in your class.

Condition 3: The second participant comes from another school.

Condition 4: The second participant comes from another class.

Each round, each of you has the choice between two options: A and B.

Your winnings are shown in the table below

(your winnings are in blue pen, your partner's in black pen)

If you choose option A, you earn 3 credits (regardless of the choice of the other participant). The other participant also receives 3 credits if he has also chosen option A. Conversely, if he has chosen option B, the other participant receives nothing, while you will get 3 credits

Please keep in mind that 1 credit that can be changed into a milk carton at your canteen (1 credit = 1 coupon)

If you choose option B and the other participant also chooses option B, you both receive 5 credits. However, if you choose option B and the other participant chooses Option A, you get nothing while the other participant will receive 3 credits.

As you can see in the Table below:

		Other Participant's Choice	
		Option A	Option B
Your Choice	Option A	3 Credit, 3 Credit	3 Credit, 3 Credit
	Option B	3 Credit, 3 Credit	5 Credit, 5 Credit

The next page will allow you to make your choices.

Once you have both chosen your option, the RAs will let you know your choice, the choice of the other participant, and the credits you are entitled to.

After this the next round will begin where you can choose between A and B again.

When you are ready, please let our the helping staff know

You have chosen option __ for this round. The other participant chose the __ option for this round. This means you have earned __ credits and the other participant __ credits.

Please continue to the next round and choose a new option.

3. Redistribution Game

In this game, you determine the credits that two other participants (A and B) will receive.

[4 variations:

Condition 1: Participant A is called[name] Participant B is called[name]

Condition 2: Participant A is in your class. Participant B is in another school.

Condition 3: Participant A is in another school. Participant B is in your class.

Condition 4: Participant A is in your class. Participant Bis in another school

Step 1

The two participants were randomly matched, then __ **credits** were allocated to participant A and __ **credit** to participant B (randomly).

Between them, the two participants therefore won 10 credits.

Please keep in mind that 1 credit that can be changed into a milk carton at your canteen (1 credit = 1 coupon)

Which distribution do you want to choose for participants A and B?

Current distribution	Your Distribution	
	Participant A	Participant B
Participant A receives 9 credits and Participant B receives 1 credit.		

Participant A receives 8 credits and Participant B receives 2 credits.		
Participant A receives 7 credits and Participant B receives 3 credits.		
Participant A receives 6 credits and Participant B receives 4 credits.		
Participant A receives 5 credits and Participant B receives 5 credits.		
Participant A receives 4 credits and Participant B receives 6 credits.		
Participant A receives 3 credits and Participant B receives 7 credits.		
Participant A receives 2 credits and Participant B receives 8 credits.		
Participant A receives 1 credit and Participant B		

receives 9 credits.		
---------------------	--	--

Use the Page now that the helping staff gives you.

4. Competition Game

In this game, our staff will give a piece of paper with a scale, you have to draw a line in the middle of a horizontal line going from 0 to 100. Our staff member will show you how to play this game. The goal is to draw the line in the middle nearest to the number 50.

[As shown in the picture below]



The next page will contain 48 of these axes. You will have 2 minutes to correctly place the greatest number of cursors out of 50.

Each correct positioning will earn you credits and you choose between two options to receive credits.

Option A: You receive **0.2 credits** for each cursor correctly positioned on 50.

Option B: You play against a partner (selected randomly).

Condition 1: The second participant is called [name]...

Condition 2: The second participant is also in your class.

Condition 3: The second participant comes from another school.

Condition 4: The second participant comes from another class.

If the number of lines that you position correctly is **greater** to the number of the other participant, you receive **0.5 credits** for each correctly drawn line

If the number of lines you have drawn correctly is **less** as the number of the other participant, **you receive nothing**.

If you draw the **same number** of lines correctly, you receive 0.2 credits for each correctly drawn line.

Which option do you prefer to receive the credits?

Option A: 0.20 credit for each correctly drawn line

Option B: 0.50 credit for each correctly drawn line, if your number is greater than the number of the other participant. If your number is lower, you get nothing.”

Please keep in mind that 1 credit that can be changed into a milk carton at your canteen (1 credit = 1 coupon)

Please draw a line nearest to the center of the scale drawn on the sheet

as soon as possible.

You have 2 minutes!

The helping staff will give you the final scores.

You have correctly drawn ____ lines.

You have chosen option A, so you receive 0.20 credits per correctly drawn line, or ____ credits in total.

You have chosen option B (game against another participant). This other participant was

able to correctly draw more lines than you so you don't get anything.

This other participant correctly drew fewer lines than you so you receive ___ credits.

Appendix D8. Multiple Hypothesis Testing

In our paper, we address the challenges of multiple hypothesis testing by implementing two statistical procedures to control the False Discovery Rate (FDR) and the Familywise Error Rate (FWER). Specifically, we utilize Anderson's (2008) two-stage method, which refines the Benjamini-Hochberg procedure for FDR control by adjusting the significance level in a sequential manner to more effectively mitigate the risk of false discoveries. For FWER, we employ the `rwolf2` package to implement the Romano-Wolf step-down procedure, a resampling-based approach that accounts for correlations among tests. Our application of these techniques does not significantly alter our main conclusions. We detail below the two methods we employ, along with the associated equations and the procedural steps that facilitate their implementation.

- **FDR:** The False Discovery Rate (FDR) represents the expected proportion of erroneous rejections of the null hypothesis among all rejections. This method is less conservative than the Familywise Error Rate (FWER), and offers greater power to uncover true positives, a feature that becomes increasingly beneficial as the number of hypotheses grows. The FDR approach is particularly useful in contexts where researchers are dealing with multiple testing scenarios, as it balances the need to avoid Type I errors with the desire to not miss out on potentially significant findings.
- **FWER:** The Familywise Error Rate (FWER) is the probability of committing at least one Type I error across all hypotheses tested. Traditional methods for controlling FWER, such as the Bonferroni correction, are highly conservative. They prioritize the minimization of false positives to such an extent that the likelihood of any individual false discovery is kept exceedingly low. However, this conservatism comes with a trade-off, as it often results in a substantial loss of power to detect true positives. The FWER is particularly relevant in studies where the cost of a Type I error is high, and the integrity of each individual test is important.

Equations and Procedures are detailed below for both methods we implemented in the paper.

- **Benjamini-Hochberg Procedure for FDR:**

1. Sort p-values: $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(M)}$
2. Find the largest k such that: $p_{(k)} \leq \frac{k\alpha}{M}$
3. Reject for all H_i for $i \leq k$

- **Romano-Wolf Stepdown Procedure for FWER:**

$$p_{adj}(i) = \max_{j \geq i} \left(\frac{\text{rank}(p_j)}{m+1} \right)$$

This method resamples the data (e.g., through bootstrap) to account for the correlation between tests, adjusting p-values to control the FWER.

Implementation in the Paper. In our paper, we utilize these methods as follows:

FDR: To effectively manage the False Discovery Rate (FDR), our paper employs Anderson's (2008) two-stage method, an enhancement of the Benjamini-Hochberg procedure. This refined approach involves a sequential adjustment of the significance level, thereby providing a more nuanced control of the FDR. The first stage determines an initial set of p-values, which are then subjected to the Benjamini-Hochberg threshold in the second stage, ensuring that the proportion of false positives remains controlled under the desired level.

FWER: In addressing the Familywise Error Rate (FWER), we utilize the `rwolf2` software package, which facilitates the implementation of the Romano-Wolf stepdown procedure. This resampling-based method is designed to adjust for inter-test correlations, by employing a resampling mechanism to recalibrate the p-values. This procedure iteratively refines the significance thresholds, thereby maintaining the probability of incurring one or more Type I errors across a family of hypotheses within acceptable bounds.

Tables 1 and 2: The p-values in Tables 1 and 2 are adjusted for multiple hypotheses. These tables include nine outcomes and four treatments, evaluated on the same sample of 607 teachers, resulting in 36 hypotheses. We applied q-value False Discovery Rate (FDR) correction and Romano-Wolf Family-Wise Error Rate (FWER) correction to the following outcomes:

- Average Gender Rights
- Economic Rights
- Political Rights
- Social Rights
- Legal Rights
- Petitions to Dowry without controls
- Petitions to Dowry with controls
- Petitions to Polygamy without controls
- Petitions to Polygamy with controls

The treatments are:

- Visual Narrative (Movie)
- Joint Movie-Curriculum
- Utilitarian
- Malleability

The robustness of our main results is confirmed through these Multiple Hypotheses Testing (MHT) corrections.

Table 3: For Table 3, we included four outcomes within the family (SAT without controls, SAT with controls, Math test without controls, Math test with controls) across the same four treatments, resulting in 16 hypotheses. Since results from other tables are not comparable with those in Table 3, they are not included in the same family of hypotheses.

Table 4: The MHT for Table 4 includes different specifications of the Math Test Score for various sub-samples, with or without interaction terms. This results in five specifications of the Math Test Score on four treatments, totaling 20 hypotheses.

Table 5: Table 5 focuses on the sub-sample of students playing against the opposite gender. The MHT family here includes four outcomes (Redistribution, Competitiveness, Cooperation, Coordination) across the same four treatments, amounting to 16 hypotheses.

Appendix: For every table in the Appendix, the MHT family includes only the tests within one table. The number of hypotheses is mentioned in the table notes.