Essays on Experimental Evidence for Improving Higher Education in Developing Countries

HARVARD
Kenneth C. Griffin

GRADUATE SCHOOL
OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Education

have examined a dissertation entitled

Essays on Experimental Evidence for Improving Higher Education in Developing Countries

presented by Ashutosh Bhuradia

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature P. Bla

Typed name: Prof. Peter Blair

Signature

Typed name: Prof. Susan Dynarski

Signature Nishith Prakash

Typed name: Prof. Nishith Prakash

Signature

Typed name: Prof. Michela Carlana

Signature ______________________

Typed name: Prof.

Date: April 14, 2026

Essays on Experimental Evidence for Improving Higher Education in Developing Countries

A dissertation presented

by

Ashutosh Bhuradia

to

Harvard Graduate School of Education

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Education

Harvard University

Cambridge, Massachusetts

April 2026

Dissertation Advisor: Professor Peter Blair | Ashutosh Bhuradia

Essays on Experimental Evidence for Improving Higher Education in Developing Countries

Abstract

This dissertation consists of three essays that present experimental evidence on improving higher education in developing countries.

The first chapter, "How Leaders Emerge," examines the interplay of gender, teamwork, and leadership through a 2x2 randomized field experiment involving 203 mixed-gender teams in a project-based competition at an engineering college in rural India. I find that female-majority teams that choose their own leaders outperform other groups by 0.38-0.51σ, driven by greater teamwork and more effective leadership.

The second chapter, "Misperceptions about Caste," studies how misperceptions about caste disparities influence attitudes toward caste and support for redistribution through affirmative action in higher education. I surveyed 774 college-aged respondents and find that while upper- and lower-caste youth both underestimate caste disparities, upper-caste youth underestimate disparities to a larger extent. Correcting misperceptions through an information intervention improves attitudes toward lower-caste groups by 0.13σ but does not alter support for affirmative action.

The third chapter, "Women Engineers," conducts a descriptive analysis of a program designed to address barriers to women's STEM education in India. Across five cohorts, the program attracts a broad applicant pool with 55-60% from rural areas and 60- 85% from low-income families. I also find evidence of a "perception gap" in confidence between low and high-income women engineers: while self-reported confidence is identical across income groups, interviewers perceive low-income women as less confident.

Contents

List of Tables

List of Figures

Acknowledgments

I am deeply grateful to my dissertation committee–Peter Blair, Michela Carlana, Susan Dynarski, and Nishith Prakash–for their mentorship and for pushing me to be more creative (and ambitious) with my research. I am also grateful to other faculty members at HGSE (and elsewhere)–Eric Taylor, Alejandro Ganimian, and Emmerich Davies–who have often gone above and beyond in giving feedback and sharing opportunities even though they do not serve a formal role on my dissertation committee.

I would also like to thank my wife, Nirali Desai, who has been a true pillar of support (as the cliche goes). Never easy being the partner of a PhD, and she has done an incredible job. I am also indebted to many colleagues, collaborators, and friends, without whom this journey would not have been possible: Saloni, Alex, Arkadijis, Julian, Lily, Sandra, Santi, Devika, Abhinav, and Lauren.

Finally, I would like to acknowledge the support of the grants and fellowships that have enabled various aspects of this dissertation: HGSE Doctoral Research Grant, the Weiss Fund, Digital Harbor Foundation Fast Grant, and the James M. and Cathleen D. Stone Program in Wealth Distribution, Inequality, and Social Policy at Harvard Kennedy School.

To my parents

Introduction

My dissertation focuses on how education in developing countries can better prepare young people for work and citizenship. My agenda weaves together two strands that meet in the classroom, the labor market, and society: building the skills that modern, team-based workplaces reward, and building mindsets students need to address societal challenges. I pursue this agenda through three experimental studies embedded in higher education settings in India.

The first strand of my research involves understanding how programs in developing countries can help students—especially women in STEM—develop the skills that contribute to better labor market outcomes. Chapters 1 and 3 of my dissertation focus on this strand. A second strand of my dissertation tests interventions that are important to students' participation in civic life and to addressing societal challenges related to inequality. This is covered in Chapter 3.

Overview of Chapters

Chapter 1: How Leaders Emerge examines the interplay of gender, teamwork, and leadership in a setting where gender norms remain strong. College students entering the workforce are increasingly expected to collaborate and lead mixed-gender teams. Yet we know little about the interplay of gender, teamwork, and leadership, especially in settings that are traditionally gender segregated. This paper examines this interplay through a 2x2 randomized field experiment involving 203 mixed gender teams in a project-based competition at an engineering college in rural India. Students are first randomly assigned to male-majority or female-majority teams and further into one of two leadership conditions: leaders assigned based on a baseline measure of emotional intelligence or chosen by their own teammates. I find that female-majority teams that choose their own leaders outperform other groups by 0.38-0.51σ, driven by greater teamwork and more effective leadership. In contrast, male-majority teams that choose their own leaders have the lowest performance—driven by free-riding, coordination failures, and ineffective leadership—while teams with leaders assigned based on emotional intelligence, regardless of gender composition, fall somewhere in between. These results imply that leadership development and team performance must account for the differing dynamics across gender groups in contexts where gender norms remain strong.

Chapter 2: Misperceptions about Caste examines the role of misperceptions in shaping attitudes toward outgroups and redistributive policies. Caste remains a salient dimension of inequality in India and a key target of redistributive policies. I study how misperceptions about caste disparities influence attitudes toward caste and support for redistribution through affirmative action in higher education. I first surveyed 774 college-aged respondents, both beneficiaries (lower caste youth) and non-beneficiaries (upper caste youth) of affirmative action. I find that while upper- and lower-caste youth both underestimate caste disparities, upper-caste youth underestimate disparities to a larger extent. I then randomly assign these respondents to an online intervention that provides them with factual information about caste disparities. I find that correcting misperceptions through this information improves attitudes toward lower-caste groups by 0.13σ but does not alter support for affirmative action. The results suggest that correcting misperceptions can shift social attitudes but may be insufficient to alter preferences for redistribution.

Chapter 3: Women in STEM examines a fully-funded online training program for women engineering students in India, analyzing five cohorts of the program since 2019. Implemented nationwide by an education start-up, the program combines a women-only environment, fully online access, self-directed learning, and mentoring to address cultural, institutional, and psychological barriers faced by women in STEM. We find that the program is successful at targeting underserved populations. Despite a highly selective admissions process (with an acceptance rate of 0.8-1.4%), between 25-50% of participants come from low-income families, and over half attend non-elite institutions. Using baseline data from the most recent cohort, we also introduce a novel baseline measure of confidence based on interviewer ratings collected during the program's selection process. We find a 'perception gap' in confidence: low-income students show similar levels of self-reported confidence as higher-income peers, yet interviewers rate them as less confident. The study sheds light on the design of targeted initiatives and assessment tools aimed at fostering inclusion in STEM.

Chapter 1

How Leaders Emerge: Gender Composition, Leader Selection, and Team Performance in India[1]

1.1 Introduction

Jobs that involve non-routine analytical team tasks have expanded rapidly over the past two decades. This shift has raised the value of non-cognitive skills such as teamwork and leadership (Deming, 2017; Edin et al., 2022). As organizations increasingly rely on teams to perform complex work, individuals are expected to collaborate and lead across gender lines. Yet, in many developing countries, gender norms and accompanying social segregation persist, creating tension between the collaborative demands of modern workplaces and the realities of gendered social and educational environments. Understanding how men and women collaborate and lead in such environments is thus central to preparing millions of students for the labor market.

A key question that emerges is how team composition and leadership selection shape collaboration across gender lines in these settings. For instance, team gender composition and the process of leadership selection can influence communication, coordination, the sharing of diverse perspectives, and conflict resolution, all of which affect team performance. However, evidence on when team gender composition matters, and how it interacts with leadership selection, remains unclear. In male-dominated STEM fields, these compositional dynamics may be especially consequential for women: female-majority settings can enhance psychological safety, whereas those with a male-majority may marginalize women. Similarly, the process of leadership selection can either legitimize women's authority or activate cultural biases that undermine it. Understanding how

[1] I am extremely grateful to Peter Blair, Michela Carlana, Susan Dynarski, Nishith Prakash, Alejandro Ganimian, Saloni Gupta, Alex Bolves, and the participants in the Economics of Education Seminar at Harvard for their feedback and helpful comments. All errors are my own. Harvard IRB approval IRB25-0406. This study is registered in the AEA RCT Registry (AEARCTR-0016103).

these forces jointly operate is essential for informing how firms organize teams and how higher education programs cultivate teamwork and leadership skills in an equitable manner. Thus my paper asks the following questions: Does a team's gender composition and leadership selection process affect its performance? Does it affect teamwork, leadership quality, and the role of women in teams?

I conduct a study to test the effect of team gender composition and leadership selection on the performance of teams in a project-based competition. Additionally, I examine the extent to which gender composition and leadership selection affects teamwork, leadership, and the role of women leaders. The study involves 203 student teams at an engineering college in rural India. I partner with the college's Training and Placement Office (TPO), which is responsible for students' job-skills training.

The context of rural India is important for several reasons. Gender shapes socialization and work in rural India, through occupational and social segregation (Montes et al., 2018). This mirrors aspects of college life. For instance, there are separate hostels (dorms) for men and women, and entry into hostel buildings is restricted to students of the same gender. In classrooms, women and men are discouraged to sit with each other. While social norms and the organization of the college itself limit mixed-gender interactions, students aspire to jobs that frequently require collaboration between men and women. These conditions provide a natural setting for understanding how gender norms influence teamwork in contexts that resemble many higher education environments across developing countries.

At the start of the study, students are administered a baseline survey after which they are first randomized into mostly 3-member mixed-gender teams, which are either male or female majority. Teams are then randomized into one of two leadership selection conditions: teams are either assigned a leader based on emotional intelligence measured at the baseline or teams nominate a team member as their leader. The choice of emotional intelligence for leadership assignment is based on prior research that links emotional intelligence to team and leader performance (Weidmann and Deming, 2021; Weidmann et al., 2024).

Teams, with their externally assigned or peer-nominated leaders, then work together on a project for two weeks. The project is part of a competition with cash prizes and involves designing a Mobile App idea that addresses challenges in rural India. The competition's cash prizes are sufficiently incentivized for this sample. For instance, 66% of participants are from families with a self-reported income of less than $1,140 per year. Top performing teams can win $140 on average, which is more than a month's family income for a majority of participants. Teams in the top decile are guaranteed to win $34, which is still more than a week's family income for a majority of participants.

Submitted projects are evaluated and scored by two trained raters who are blind to the treatment conditions. Before the results from this evaluation are announced though, each participant is asked to fill out an endline survey. In the endline survey, participants rate the effectiveness of their team and team leader. They also report on their and other team members' contribution to the project. Top-20 scoring teams (the finalists) are then selected to participate in a in-person competition at the college.

I find that female-majority teams that nominate their own leaders outperform other teams. Specifically, female-majority teams that nominate their own leader score 0.51σ higher on their projects than male majority teams that nominate their own leader, 0.47σ higher than male majority teams that have leaders assigned to them, and 0.38σ higher than female-majority teams who have leaders assigned to them. Moreover, female-majority teams that nominate their own leaders are also more successful in advancing to the final round–13% of their teams are finalists, marginally (though not statistically) higher than female-majority and male-majority teams with assigned leaders at 11% and 10.5% respectively, and significantly higher than male-majority teams that nominate their own leader of whom only 4% make it to the final round. The results suggests that while female-majority teams tend to perform better on average, their overall performance gains are especially pronounced when they have agency to choose their own leader. In contrast, agency in choosing leaders does not benefit male-majority teams, and even undermines their performance at the top end of the distribution.

To understand factors that might contribute to these differences in performance, I examine outcomes related to teamwork, leadership, and contribution to project work. I find that female-majority teams consistently report better teamwork than male-majority teams. Across the teamwork measures, their scores are between 0.18σ and 0.34σ higher. As such, female-majority teams (with nominated or assigned leaders) report better teamwork than their male counterparts but there are no differences in teamwork between female-majority teams.

In contrast to teamwork, leadership outcomes paint a more nuanced picture and especially help understand the differences in performance between female-majority teams despite similar levels of teamwork. I capture three leadership outcomes. The first outcome is a leader effectiveness score on a scale of 1-10. The second is leadership quality index, a holistic measure constructed from five questions that capture different dimensions of leadership such as coordination, communication, and conflict resolution. Each team member rates their leader on these measures. Finally, I report how much team leaders are perceived to contribute to their team's work. Each team member reports the percentage of work they and other teams members (included the team leader) contributed to the project.

As before, female-majority teams that choose their own leader score 0.26σ higher on leader effectiveness and 0.42σ higher on leadership quality index compared to male-majority teams. However, unlike before, there are significant differences between female majority teams. Specifically, female majority teams that choose their own leader score 0.43σ higher on leader effectiveness, 0.33σ higher on leadership quality, and 8.34pp higher on leaders' contributions compared to female-majority team have been assigned a leader. There are no such differences between male-majority teams based on leadership structure. This suggests how leaders are selected matters more for female-majority teams than it does for teams with a male-majority. Taken together, gender composition of teams matters more for teamwork, and leadership structure matter more to leaders' effectiveness and their contribution to collaborative work, especially in female-majority teams.

Given the differences in leaders' effectiveness and quality, do the type of leaders differ across conditions? For example, in this gender segregated and patriarchal setting, one might expect women might be less likely to be chosen leaders in male-majority teams (accounting for the proportion of women) and less likely to be leaders if their peers choose leaders versus when leaders are assigned externally. However, I find that leadership selection does not affect the gender representation of leaders, and both processes (leaders externally assigned or nominated by peers) yield nearly identical shares of women leaders. Gender composition, in contrast, mechanically shapes who becomes the leader because teams draw from pools that are themselves more male or more female. For example, among teams that nominate their own leaders, 71.4% of leaders in male-majority teams (with 66.6% men) are men, and 75.5% of leaders in female-majority teams (with 66.6% women) are women. These figures indicate that once we account for the underlying gender composition of teams, leadership selection shows no statistically significant gender disparity; teams tend to select leaders from the majority gender at rates that closely mirror the pool. As a result, when aggregating across all teams, the overall share of women leaders (53%) approximates their proportion in the sample (50.8%).

What then is the source of variation in the quality and performance of leaders across conditions? I observe three main sources of variation. First is the variation that arises from the experimental design: female-majority teams are much more likely to have women leaders. And if women leaders, and women in general, are systematically different than men in the baseline, then part of the variation can be explained by these differences. I find suggestive evidence of differences by gender. While men and women have similar ability (based on baseline IQ and EQ scores), women are more likely to be low-income by 12.6pp and have a stronger preference for leadership by 0.34σ. Thus, strong pecuniary incentives combined with a higher willingness to lead might drive performance and engagement in teams with more women. However, this variation would not explain the difference in performance between female-majority teams nor the *lack of difference* in performance between male and female-majority teams with assigned leaders.

The second source of variation is emotional intelligence, which is partly mechanical since leaders in the assigned-leader condition are selected based on emotional intelligence scores. Specifically, assigned leaders have 0.47σ higher emotional intelligence compared to nominated leaders–however, leader effectiveness scores and leader contributions for teams with assigned leaders suggest that higher emotional intelligence does not necessarily translate into better leadership. The third source of variation is leaders' willingness to lead. Leaders in female majority teams have a 0.22σ higher preference for leadership compared to leaders in male majority teams, reflecting both women's higher baseline preference for leadership and the greater likelihood of women becoming leaders in these teams. Separately, nominated leaders have a 0.23σ higher preference for leadership than assigned leaders, suggesting that peer deliberation favors individuals who are more motivated to take on the role.

These patterns in higher emotional intelligence among assigned leaders and higher willingness to lead among nominated leaders may reveal an important insight: emotional intelligence does not necessarily translate into a motivation to lead. Consistent with this, emotional intelligence and preference for leadership are only very weakly correlated ($\rho = 0.08$). This suggests that while emotional intelligence may matter, a leader's motivation to lead appears to play an independently important role in driving leadership performance.

Finally, an important source of the variation is the process of leadership selection itself. Leaders elected through a consultative process might have more legitimacy than leaders who are assigned by external actors. Women often seen as less legitimate in leadership roles, unless they legitimize their roles (Vial et al., 2016). Democratic nomination of leaders could potentially serve to legitimize leaders, especially women, which could contribute to greater effectiveness of leaders in teams that nominate leaders. Whereas leaders in female-majority teams that have been assigned a leader, most of which are women, might been seen to have lower legitimacy.

Why then do male-majority teams that choose their own leaders not reap some of the legitimizing benefits of a democratic nomination process? One reason could be than men might be harsher in dealing with other men, especially when competing for leadership roles. Several studies indicate male to male negotiations can lead to less cooperation and worse outcome in competitive settings (Sutter et al., 2009; Eckel et al., 2021; Castillo et al., 2013) compared to female to female and male to female negotiations. This could be the case in male-majority team that choose their own leader–men negotiate to assume leadership and those who are not nominated as leaders might be reluctant to cooperate and contribute to the project. I observe some evidence of this non-cooperation through free-riding–specifically, more free-riding in male-majority teams that choose their own leaders, driven entirely by men in the team who are not nominated leaders.

This study contributes to several strands of literature. First, it adds to our understanding of teamwork especially in the context of mixed-gender teams, which are increasingly common in educational and workplace settings. Gender composition of teams affects business decisions, scientific output, and performance (Yang et al., 2022; Apesteguia et al., 2012a; Truffa and Wong, 2025). Evidence of the effect of team's gender composition on performance, in particular, is mixed or unclear. While Hoogendoorn et al. (2013), Berge et al. (2016) and Lamiraud and Vranceanu (2018) find that gender-equal or female-majority teams perform better, Apesteguia et al. (2012b), Aparicio Fenoll and Zaccagni (2022), and Hardt et al. (2025) find that male-majority or all-male teams do better. Even less clear is how the evidence would apply to a real-world setting. Lamiraud and Vranceanu (2018) and Berge et al. (2016) conduct their experimental study in a lab, and Hardt et al. (2025) and Aparicio Fenoll and Zaccagni (2022) focus on narrowly defined tasks (e.g., solving math problems in teams) that are less akin to non-routine cognitive team tasks that individuals might collaborate on in a workplace. In Apesteguia et al. (2012b), the teams are endogenously formed, thus limiting causal interpretation. While Hoogendoorn et al. (2013) conduct a field experiment, they are unable to identify mechanisms that drive performance between teams of different gender composition. My field experiment contributes to this literature by incorporating a hands-on project that closely resembles a non-routine cognitive team task and tests a number of mechanisms to better understand the determinants of performance.

Second, I contribute to a large and evolving literature on leadership selection. Experimental studies in both the lab and the field demonstrate that leadership selection matters to the performance of both leaders and teams teams (Deserranno et al., 2019; Englmaier et al., 2024; Levy et al., 2011; Reuben and Timko, 2018; Brandts et al., 2015; Chemin, 2021; Weidmann et al., 2024). Leaders who are elected by team members or assigned based on observable ability tend to perform better than randomly elected leaders (Levy et al., 2011; Reuben and Timko, 2018; Brandts et al., 2015; Chemin, 2021), although in a lab setting randomly assigned leaders perform better than those who nominate themselves (Weidmann et al., 2024). However, the literature is sparse on the differences in the performance of teams and leaders when one set of leaders are elected, and another are assigned based on ability. Learning more about this distinction is important because leaders are seldom randomly selected in real-world scenarios—they are often selected based on observable ability or through deliberation between stakeholders. My study shed lights on the conditions under which different modes of leadership selection enhances or inhibits performance in the field.

Third, I contribute to the literature on leadership selection and the role of women leaders, especially in mixed-gender teams. Leadership selection processes determine the representation and legitimacy of women leaders in a variety of settings (Eckel et al., 2021; Gangadharan et al., 2016) and gender composition further affects women's willingness to become leaders and their influence within teams (Born et al., 2022; Chen and Houser, 2019; Karpowitz et al., 2024). While women leaders are more effective in certain contexts, gender stereotypes tend to lower evaluations of their leadership (De Paola et al., 2022). I see evidence of this in my study–women leaders contribute more than male leaders but are evaluated similarly or lower than their male leaders. My study thus contributes to understanding how gender composition and leadership structure interact to enhance the role, performance, and evaluation of women leaders. Finally, this paper is among the few to experimentally study teamwork and leadership in a developing-country context. By embedding a randomized design in an engineering college in India, it provides causal evidence on how gender composition and leadership selection shape collaboration and performance in settings where gender norms are still highly salient.

This paper is structured as follows. In Section 2 and Section 3, I describe the research design and the sample, and test for balance. In Section 4, I present the main results and mechanisms. In Section 5, I discuss the implication of my findings and conclude.

1.2 Research Design

I partnered with an engineering college in a predominantly rural part of Central India. The college has undergraduate programs in engineering and technology such as Computer Science, Data Science, Information Technology, Electrical Engineering, Electronics and Communications Engineering, Agriculture Engineering, etc. Within the college, I partnered with the Training and Placement Office (TPO). Training and Placement Offices are influential departments within Indian colleges and are primarily responsible for job-skills training for students and establishing partnerships with companies to facilitate the hiring of students from the college. The TPO's interest in this study arose from feedback it received from recruiters about the importance of soft-skills (teamwork, leadership, communication skills, etc.) in hiring and deficiencies in the college coursework in developing these skills.

I enrolled 610 students to participate in the study (380 Year-1, 166 Year-2, and 64 Year-3 students). As part of the recruitment, students were told that they would participate in a team-based competition related to engineering and technology, but they were not told the specifics of the competition–who they will team-up with, what specific activity they would have to perform, and what the prizes were. Students who signed-up to participate in the competition were then asked to complete a baseline survey and assessment which captured students' basic demographic information, experience with teamwork and leadership, preferences for teamwork and leadership, IQ (through a shortened version of the Raven test), and their emotional intelligence. The assessment of emotional intelligence was based on PAGE – Perceived AI-Generated Emotions

– originally developed by Weidmann and Xu (2024), which I adapted for the context of India.[2]

After the baseline, I randomized students into within-year teams. I followed a two-step randomization protocol which proceeded as follows: first, I randomly assigned students in the same cohort into teams and varied whether students were assigned to female-majority (2 female, 1 male) or male-majority (2 male, 1 female) teams.[3] In total, students were assigned into 203 teams (96 male-majority and 107 female-majority teams). Next, I randomly assigned teams into one of two leadership selection conditions: in one condition, the team member with the highest PAGE score (proxy for emotional intelligence) within each team was assigned to be the team leader (assigned-leader arm with 101 teams); in the other condition, teams were asked to deliberate among themselves for a day and nominate or choose their own team leader (nominated-leader treatment arm with 102 teams).[4] In Table 1.1, I show the distribution of team leader selection arms by team gender composition.

Table 1.1: *Distribution of Leader Assignment by Team Gender Composition*

Team Gender Composition	Assigned Leader	Nominated Leader
Female-Majority	54	53
Male-Majority	47	49

Notes: The table shows the number of teams by gender composition and whether the team leader was assigned or nominated. Leader assignment was stratified by team gender composition.

After team-formation but before informing teams of their leadership selection treatment assignment, teams were brought together in classrooms for a brief session. With the exception of team leader assignment, the sessions were identical but separate for teams in the leader-assigned and leader-nominated conditions. The teams did not know they were allocated to these two different conditions. During this session, team members met and introduced themselves, and played an ice-breaker game. The teams were then told more details about the competition in

[2]Details on test development and the psychometric properties of PAGE are provided in the Appendix, Section C.

[3]Because numbers within a year were not always perfectly divisible by 3, two teams had 4 members each and one team had 2 members – one team with 2 female and 2 male, one team with 3 female and 1 male, and another with 1 female and 1 male member. For consistency, I code both gender balanced teams as female-majority. In the Robustness Checks section, I present results from additional analysis to show that dropping these teams altogether from my analysis does not alter the main findings.

[4]Thirteen teams in the assigned-leader treatment arm had 2 members with the same PAGE score. In that case, one of the team members was randomly assigned as the team leader.

which they would participate. The competition involved working in their teams to develop a Mobile App idea to solve a problem in rural India. They were provided a clear criterion on which their projects would be judged (project scoring rubric in Appendix C3). They were told that the competition involved 203 teams and that they had 2 weeks to work on and submit their projects. In addition, they were told that the top 20 projects from a pool of 203 would be shortlisted to compete in an in-person competition at the college. Teams in the top 20 were also guaranteed cash prizes of Rs 3000 ($34.86) per team. In addition, there were separate prizes for the in-person winners. The top ranked team would win Rs 15,000 ($174.32), the second-ranked team would win Rs 12,000 ($139.45), and the third-ranked team would win Rs 10,000 ($116.21). At the end of the session, teams in the assigned leader arm were told the name of the assigned leader, and teams in the nominated-leader arm were told to deliberate and nominate their leader the next day and to simultaneously start working on their projects.

After the session, I shared Google Slides templates with each team. The templates were shared with each individual student's email accounts, which they provided in the baseline. I also collected information on the identity of the team leaders in the nominated-leader treatment arm via a Qualtrics link sent to each team the following day. Student teams worked on their projects for 2 weeks and submitted the project either via the shared Google Slide, or as separate pdf or presentation sent to the research team via email. Of the 203 teams, 156 (76.3%) submitted a project.

After the project submission deadline, I reached out individually to the students to complete an endline survey. The survey elicited students' perception of their teams' and team leaders' effectiveness, perceptions of their own and their teammates' contribution. During the same time, I hired and trained two evaluators to score each project. The raters were blinded to the treatment conditions. Specifically, raters were blinded to the gender composition (names were removed from submissions before the rating process) and the leadership structure of the teams. After each project had been scored, I ranked the teams based on their total score averaged across two evaluators. The top-20 ranked teams were announced on the WhatsApp communities that consisted of all teams. Each of the top 20 teams was also separately informed and invited to compete in the final round. The final in-person round was held in the college auditorium. Each team presented for 10 minutes in a randomly assigned order. A panel of outside judges with a background in rural development and technology relevant to the competition was invited to assess the presentations.

The winners were announced on the same day and the cash prizes were distributed in person to each team (see Appendix B for images from the competition). Figure 1.1 summarizes the timeline for the implementation and data collection activities described above.

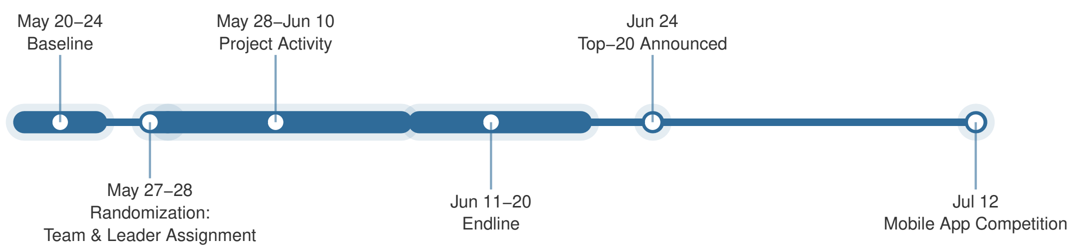


Figure 1.1: *Implementation Timeline, May-July 2025*

1.3 Sample and Balance Test

As reported in Table 1.2, a majority of the participants are low-income (self-reported annual family income less than Rs 1 lakh or $1140) and lower-caste. Compared to the male participants, female participants comprise 50.8% of the sample, are 12.6 percentage points more likely to be low-income (p-value = .001), and have a 0.34 SD (p-value = .000) higher preference for leadership.[5] However, I do not find differences in ability (IQ as measured by Raven, and emotional intelligence as measured by PAGE) between men and women in the sample.

I further test the correlation between ability, gender, and preference for leadership. I show these in Figure 1.2. I find a moderate positive correlation (0.37) between Raven and PAGE (emotional intelligence and IQ), and a positive weak correlation (0.18) between Raven and preference for leadership. Although I find a negligible correlation between gender and ability, I observe a positive, if weak, correlation (0.17) between being female and the preference for leadership. The positive

[5]To measure preference for leadership, I asked students to indicate how much they'd like to be a Team Leader in the upcoming competition on a scale of 1 (I really don't want to be a team leader) to 10 (I really do want to be the team leader). Furthermore, the question said that 'team leaders are responsible for directing the group and making final decisions. The role will involve communicating with your team members, delegating, collating information, and making decisions. Please be as honest as possible.'

Table 1.2: *Differences in Baseline Characteristics by Gender*

Variable	**Mean (All)**	**Mean (Male)**	**Mean (Female)**	**Difference**	***p*-value**
Low income	0.667	0.603	0.729	0.126^{***}	0.001
Upper caste	0.284	0.299	0.269	-0.030	0.427
PAGE score (std.)	0	-0.029	0.028	0.057	0.490
Raven score (std.)	0	0.067	-0.065	-0.132	0.103
Team lead pref. (std.)	0	-0.174	0.168	0.342^{***}	0.000
Percent of sample		*49.2%*	*50.8%*		

Notes: The table reports mean values for male and female students and the p-value from a two-sided t-test of the difference in means. Stars denote statistical significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. PAGE score, Raven score, and Team lead preference are standardized to have a mean of 0 and standard deviation of 1.

relationship between being female and the preference for leadership is in contrast to Western settings, where the preferences for leadership are positively correlated with being male(Weidmann et al., 2024).

Figure 1.2: *Correlation between Gender, Ability, and Leadership Preference*

Notes: *Pairwise Pearson correlations among gender, and standardized measures of emotional intelligence (PAGE), IQ (Raven), and leadership preference. Diagonal omitted.*

I conduct a two-stage randomization to first randomize individuals within academic years into male and female majority teams, and then the teams into assigned-leader and nominated-leader arms. Thus, I conduct two balance tests. First, following Karpowitz et al. (2024), I test for within-gender differences between male and female students assigned to teams of different gender composition. I report this balance in Table 1.A1. I find no statistically significant within-gender differences between individuals assigned to male and female-majority teams. Second, I test for balance between teams in the assigned and nominated leadership arms. I report this balance in Table 1.A2. Here, too, I find no statistically significant differences between teams assigned to different leadership arms. However, to demonstrate the robustness of my findings, I report both covariate-adjusted results in section 4.7 (*Robustness Checks*).

1.4 Results

1.4.1 *Effect of Team Gender Composition and Leadership on Performance*

In this section, I report results on team performance, teamwork, and leadership. The main outcome of interest is how well teams perform on their projects. As noted earlier, teams worked for two weeks developing App ideas that address challenges in rural India. The projects focused on challenges related to agriculture or education. The projects were then scored by two trained evaluators. The scoring was based on a pre-defined rubric shared with students at the beginning of the competition. The projects were evaluated according to the following criteria: problem statement, feasibility of the solution, user interface quality, technical innovation, and potential social impact. For each criterion, a team could score a minimum of 0 and a maximum of 4 points. All criteria were equally weighted. The projects that were not submitted received a score of 0. Thus, overall performance is a cumulative measure of whether teams submitted the projects and scores of teams that did.

I use the following regression specification to examine the effect of team gender composition and leadership on performance:

$$\begin{aligned} Y_i = \beta_0 \ + \ & \beta_1 \, \text{FemaleMajority}_i + \beta_2 \, \text{Assigned}_i \\ + \ & \beta_3 (\text{FemaleMajority}_i \times \text{Assigned}_i) + \delta_{\text{Year}(i)} + \varepsilon_i \end{aligned} \tag{1.1}$$

where Y_i denotes the outcome for team i. FemaleMajority_i is an indicator variable equal to one if the team is female-majority and zero if the team is male-majority, Assigned_i is an indicator equal to one if the team leader was assigned based on PAGE scores (emotional intelligence) and zero if the leader was nominated, and $\text{FemaleMajority}_i \times \text{Assigned}_i$ is the interaction between the two indicators. Because individuals are randomized into teams within academic cohorts, I also include year (cohort) fixed-effects, $\delta_{\text{Year}(i)}$. ε_i is the error term at the team level.

However, to aid interpretation, I estimate and present the following pairwise comparisons: within each leadership structure (nominated or assigned leader), I show the difference between teams of female and male-majority; then within each gender composition (male or female-majority), I show the differences between teams with a nominated versus those with an assigned leader. These pairwise comparisons correspond directly to the four experimental arms reported in Table 1.1. I report results for the four main team performance outcomes: (1) project completion; (2) project scores for teams that completed; (3) total project score (assigning zero to non-completers); and (4) the probability of being a finalist.

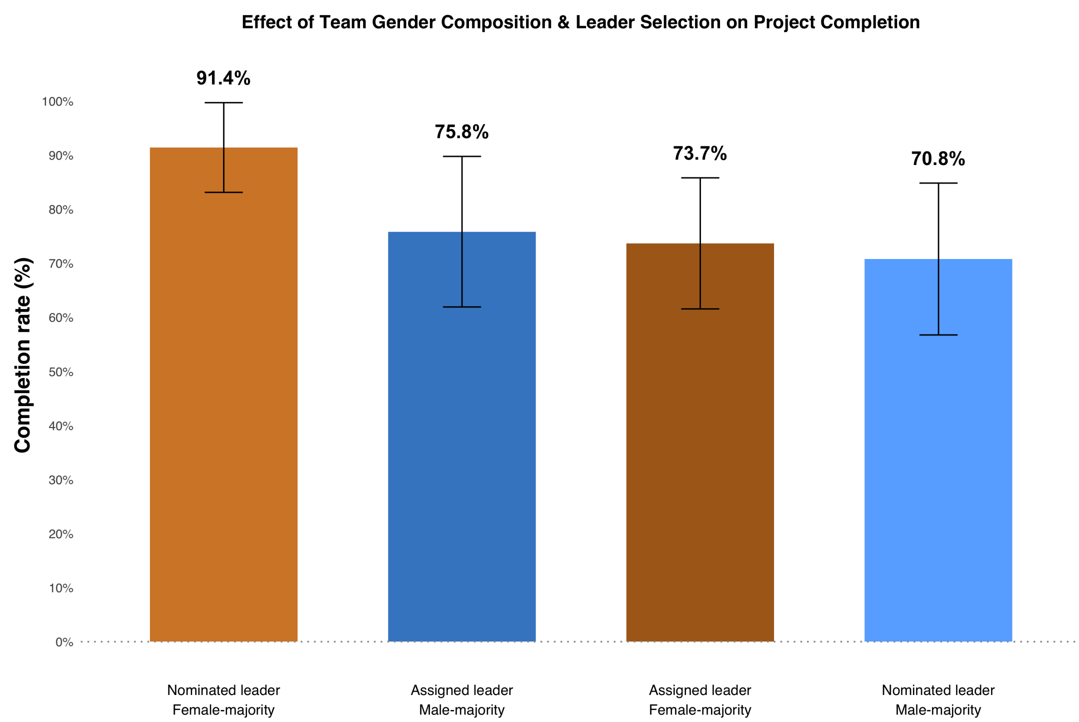


Figure 1.3: *Effect on Gender Composition & Leadership Selection on Completion Rate*

Notes: *This figure shows project completion rates for each treatment group. The outcome is a binary variable, 'complete', which indicates whether a team completed (and submitted) their project or not. Each bar shows the mean of the outcome for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

In Figure 1.3, I show completion rates for all four groups. Female-majority teams with *nominated* leaders have the highest completion rate (91.4%), followed by male-majority teams with *assigned* leaders (75.8%), female-majority teams with *assigned* leaders (73.7%) and male-majority teams with *nominated* leaders (70.8%).

Column (1) of Table 1.3 tests whether these differences in completion rates are statistically significant. First, I compare teams within the same leadership structure (nominated or assigned). Among teams with nominated leaders, female-majority teams are 20.6 percentage points more likely to complete than their male-majority counterparts ($p < 0.01$). Among teams with assigned leaders, the difference in submission rates between female- and male-majority teams is small and statistically insignificant.

Next, I compare teams with the same gender composition (female or male-majority). Among female-majority teams, assigning (rather than nominating) a leader reduces the completion rate by 17.7 percentage points ($p < 0.05$). Among male-majority teams, the roughly 5 pp higher completion rate under assigned versus nominated leadership (75.8% vs. 70.8%) is not statistically significant. Taken together, the completion rate is highest when female-majority teams nominate their own leader, and external assignment largely eliminates that advantage. However, the process of leadership selection does not affect completion rates for male-majority teams. These results indicate that the combination of female-majority composition and leader nomination yields the highest completion rates, while external assignment dampens that advantage.

Table 1.3: *Performance by Gender Composition and Leadership Structure*

	Completed	Project score (z)	Project score (z), 0 for incomplete	Finalists (%)
Panel A: Within Leadership Structure				
Nominated leader: Female-majority (=1)	0.206***	0.171	0.502***	0.123**
	(0.080)	(0.221)	(0.190)	(0.056)
Assigned leader: Female-majority (=1)	-0.022	0.277	0.082	0.036
	(0.089)	(0.261)	(0.212)	(0.065)
Panel B: Within Gender Composition				
Female-majority teams: Assigned (=1)	-0.177**	-0.045	-0.382**	-0.023
	(0.073)	(0.181)	(0.174)	(0.062)
Male-majority teams: Assigned (=1)	0.051	-0.150	0.038	0.065
	(0.092)	(0.278)	(0.216)	(0.053)
Observations (teams)	203	156	203	203

Notes: Each cell reports the coefficient on the listed indicator from a separate OLS regression with year fixed effects. Heteroskedasticity-robust standard errors, clustered at the team level, are in parentheses. Intercepts omitted. "Completed" is an indicator for whether a team completed (and submitted) the project or not. "Project score (z)" are standardized project scores for the subsample of completers. "Project score (z), 0 for incomplete" are standardized scores that include all teams–and teams with incomplete projects are assigned a score of 0. "Finalists" is an indicator for Top-20 teams. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The second outcome is project scores for teams that completed (and subsequently submitted) their projects. Figure 1.4 shows standardized project scores for the completers. The figure shows that female-majority teams perform slightly better than male-majority teams, although none of these differences are statistically significant. Column (2) of Table 1.3 tests these differences and confirms the pattern displayed in Figure 1.4.

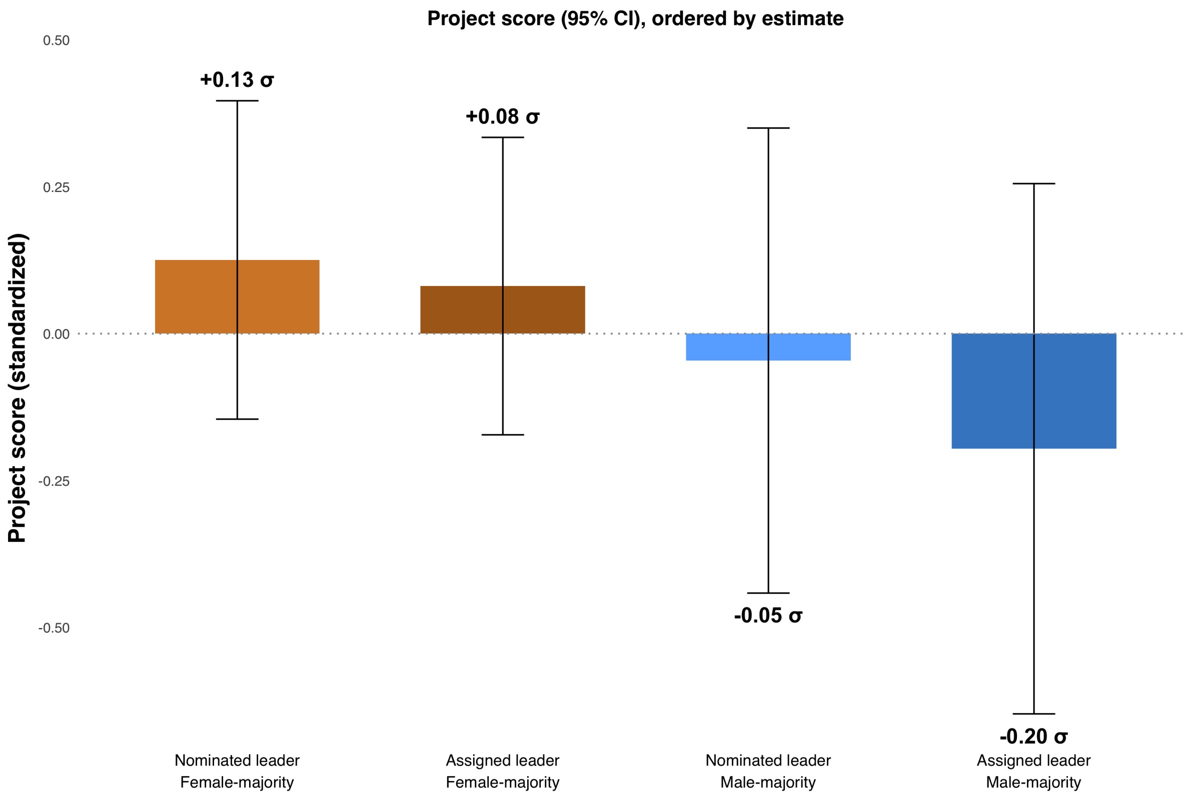


Figure 1.4: *Effect on Gender Composition & Leadership Selection on Project Scores*

Notes: *This figure shows project scores for each treatment group. Scores are estimated at the team-level. Each team's project is scored by two trained raters. The project scores are then calculated by averaging the two scores and then by standardizing them to have a have mean of zero and a standard deviation of one. Teams that did not complete the project are excluded from this analysis. Thus, this figure compares the subsample of 'completers'. Each bar shows the mean of the outcome. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

I first compare teams within the same leadership structure (nominated or assigned). Among teams with nominated leaders, female-majority teams score about 0.17σ higher than male-majority teams; among teams with assigned leaders, the female–male majority difference is approximately 0.27σ—both differences are large but imprecisely estimated. Comparing teams with the same gender composition, nominating versus assigning a leader in female-majority teams leads to a modest but statistically insignificant increase of about 0.05σ; among male-majority teams, the analogous difference is roughly 0.15σ and likewise insignificant. Conditional on completion, project quality appears comparable across all groups.

To capture overall team performance, I assign zero to teams that did not submit and standardize total scores across all teams. This outcome shows the combined effect of completion and project quality. Figure 1.5 shows that female-majority teams with nominated leaders substantially outperform all other groups, with a mean of roughly $+0.36\sigma$, compared to near-zero or negative scores for others. Column (3) of Table 1.3 quantifies these patterns.

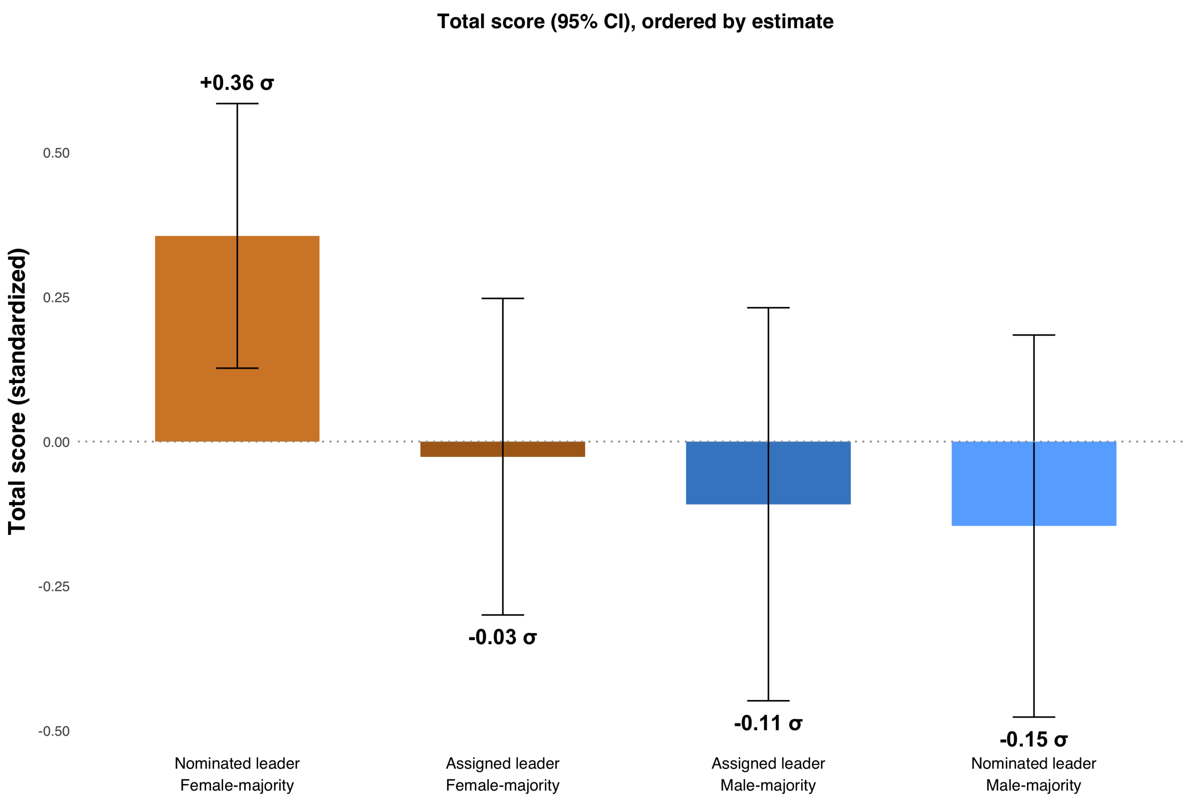


Figure 1.5: *Effect on Gender Composition & Leadership Selection on Total Scores*

Notes: *This figure shows project scores each treatment group. Outcome is a z-score, standardized to have a mean of 0 and a standard deviation of 1. This includes all teams, with teams that did not complete (and submit) their projects assigned a score of 0. Each bar shows the mean score for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

When comparing teams within the same leadership structure, female-majority teams with nominated leaders outperform their male-majority counterparts by 0.51σ. Among teams with assigned leaders, the gap between female and male-majority teams is small (0.08σ) and not significant. For teams with the same gender composition, female-majority teams that nominate their own leader score 0.38σ ($p < 0.05$) higher than female-majority teams that are assigned a leader. For male-majority teams, leadership assignment makes little difference. Overall, female-majority teams that nominate their own leader outperform all other teams by a significant margin. This pattern suggests that agency in leadership selection benefits female-majority teams, while it does not benefit male-majority teams.

Finally, Figure 1.6 presents the share of teams that reached the finalist stage. For interpretability, the figure reports raw finalist rates (without year-fixed effects), while Column (4) of Table 1.3 shows estimates with year fixed effects. Female-majority teams with nominated leaders have the highest finalist representation (13.1%), followed by female-majority teams with assigned leaders (11.1%), male-majority teams with assigned leaders (10.6%), and male-majority teams with nominated leaders (4.1%).

Within leadership structure: among teams with nominated leaders, female-majority teams are 12.3 percentage points more likely to reach the finalist stage than their male-majority counterparts ($p < 0.05$). Among teams with assigned leaders, the female–male difference is smaller (about 3.6 pp) and statistically insignificant. *Within gender composition:* for female-majority teams, assigning rather than nominating a leader slightly reduces the finalist rate by about 2.3 pp; for male-majority teams, assigned leaders lead to 6.5 pp higher finalist rate. However, neither of these estimates is statistically significant.

In summary, the performance of female-majority teams that nominate their own leaders seems to be driven by 'getting things done', as reflected by these teams completing their projects and getting them over the finish line. Even when accounting for only completed projects, teams with female-majority that nominate their leader score the highest on average (0.13σ) and reach the final round at a higher rate (13%). Taken together, the results suggest that performance gains of female-majority teams are amplified when they have the autonomy to choose their own leaders. In contrast, agency in leadership selection does not improve outcomes for male-majority teams and may even dampen their performance at the top of the distribution.

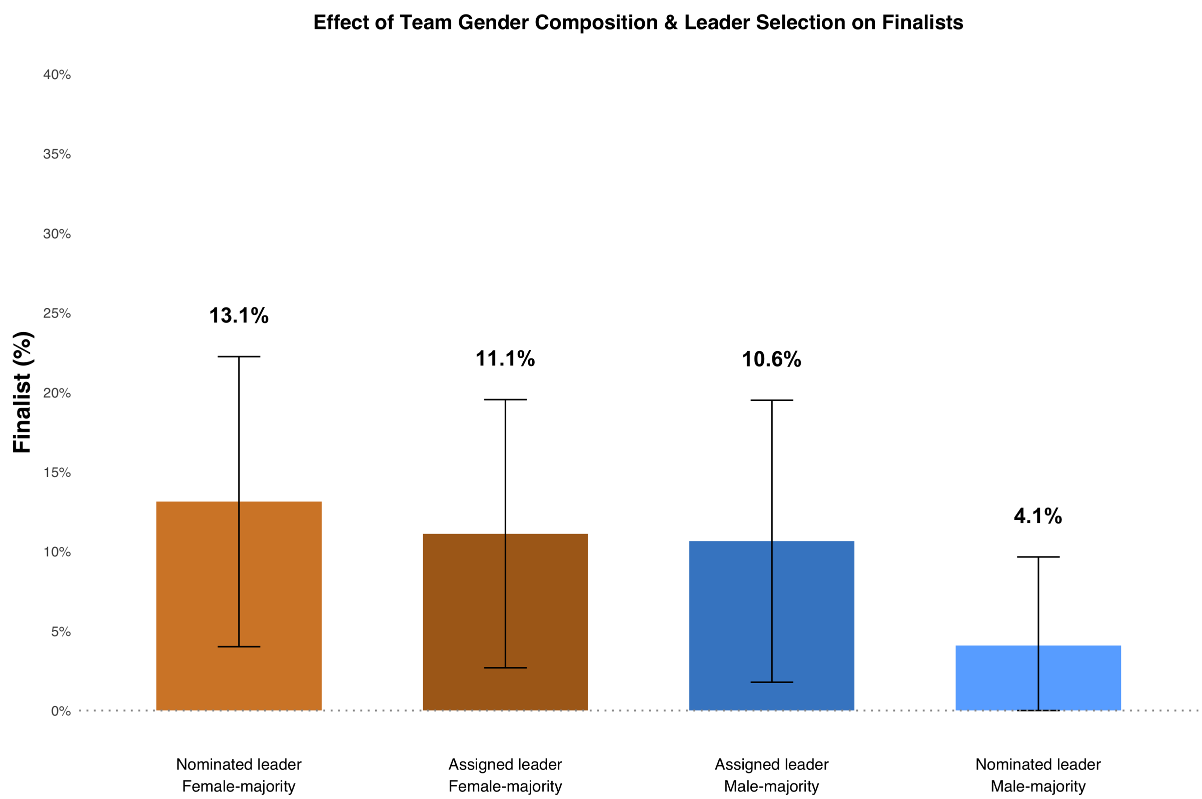


Figure 1.6: *Representation of Teams in the Final Round*

Notes: *This figure shows the share of teams within each groups that were finalists. Bars represent the raw proportion of teams within each groups that reached the final round, with percentages displayed above each bar. The y-axis is expressed as the percentage of teams from that group that made it to the final round.*

1.4.2 *Effect on Teamwork*

To understand why female-majority teams—particularly those with nominated leaders—perform better, I examine several underlying components, including teamwork, and leaders' effectiveness and contributions. I begin by analyzing teamwork, using two complementary measures: one for team effectiveness and another for team equity. I construct these two indices that capture team effectiveness and team equity, based on endline survey items, which are standardized using factor analysis.[6] Both indices load cleanly onto single factors and exhibit high

[6]The Team Effectiveness Index was constructed from five Likert-scale items (e.g., "Our team worked together effectively") that capture overall team functioning and performance. The Team Equity Index reflects students' perceptions of fairness, voice, and inclusion within the team and was derived from six items (e.g., "Team members from all backgrounds were equally respected") All items used a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree).

internal consistency (Cronbach's α = 0.92 for effectiveness, 0.80 for equity).[7]

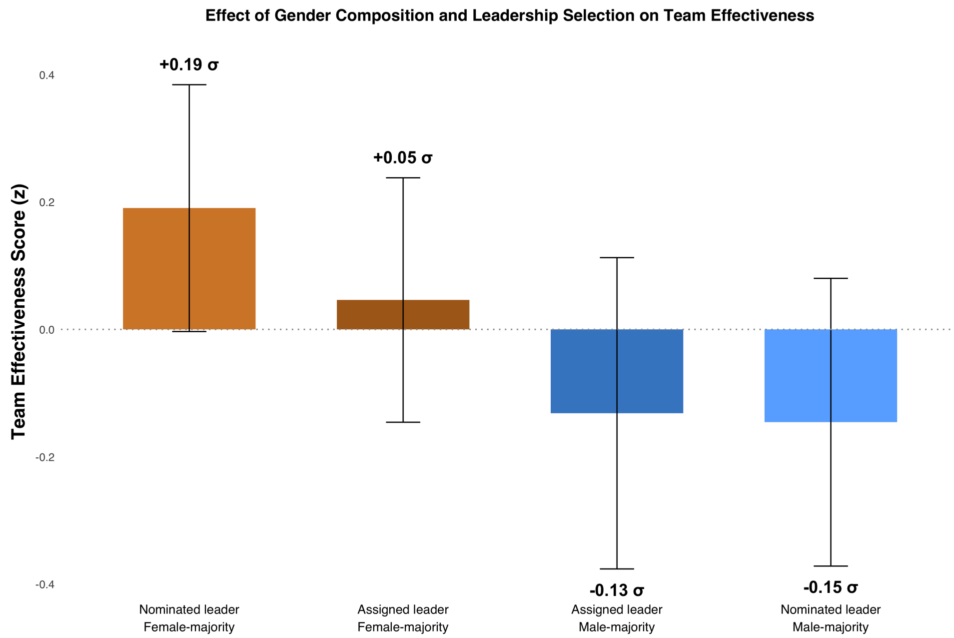


Figure 1.7: *Effect on Gender Composition & Leadership Selection on Team Effectiveness*

Notes: *This figure shows team effectiveness scores for each treatment group. Team effectiveness index is constructed using exploratory factor analysis from responses to five survey items, and standardized to have a have mean of zero and a standard deviation of one. Each bar shows the mean of team effectiveness score for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

Figures 1.7 and 1.8 plot standardized team outcomes for each of the treatment groups. Table 1.4 tests whether differences across groups are statistically significant. I find female-majority teams with nominated leaders report substantially stronger teamwork experiences than their male-majority counterparts: team effectiveness is 0.34σ ($p < 0.05$) and team equity is 0.32σ ($p < 0.05$) higher. When leaders are externally assigned, differences remain positive and meaningful—female-majority teams report 0.18σ deviations higher team effectiveness (though not statistically significant) and 0.23σ higher team equity (p-value < 0.1) compared to male-majority teams.

The bottom panels of Table 1.4 compare teams with the same gender composition. Among female-majority teams, assigning a leader slightly reduces teamwork: team effectiveness falls by about 0.14σ and team equity by 0.12σ, though neither difference is statistically significant. Among

[7] A pooled PCA of all 11 items shows the constructs are correlated ($r = 0.59$) but load distinctly on two factors, supporting the decision to treat team effectiveness and equity as related but conceptually separate outcomes.

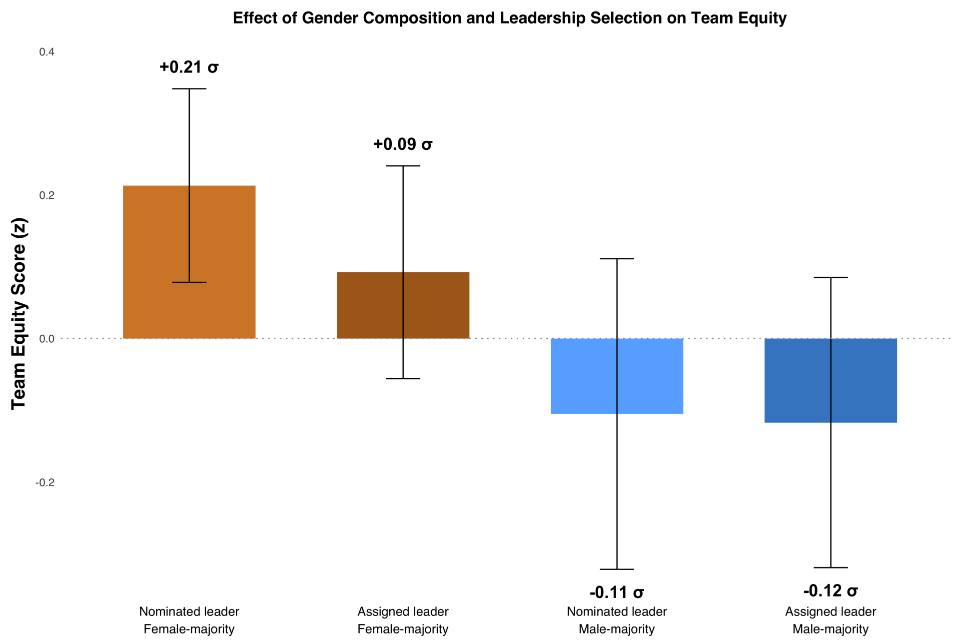


Figure 1.8: *Effect on Gender Composition & Leadership Selection on Equitable Teamwork*

Notes: *This figure shows team equity scores for each treatment group. Team equity index is constructed using exploratory factor analysis from responses to six survey items, and standardized to have a have mean of zero and a standard deviation of one. Each bar shows the mean of team equity score for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

male-majority teams, assigning a leader does not materially change teamwork, with near-zero differences in both indices.

Taken together, Figures 1.7 and 1.8 show that gender composition—particularly being on a female-majority team—plays a larger role in shaping teamwork quality than how leaders are selected. Female-majority teams consistently report stronger collaboration and a greater sense of fairness within teams, whether leaders are nominated or assigned.

1.4.3 *Effect on Leader Effectiveness, Leadership Quality, and Contributions*

Although perceptions of teamwork could be a mechanism that explain the differences in performance between female and male-majority teams that nominate their own leaders, they do not explain differences in performance between female-majority teams. Female-majority teams, regardless of their leadership structure, show similar levels of teamwork. The differences in performance between teams with a female-majority remain unexplained. One hypothesis could be that since the primary difference between female-majority teams is their leadership structure, measures of leader effectiveness and quality could shed light on these differences and explain the variation in performance between female-majority teams. Thus, in Figures 11-13, I test whether there are differences in the effectiveness, quality, and contributions of team leaders–as perceived by their peers–across treatment arms, and especially between female-majority teams.

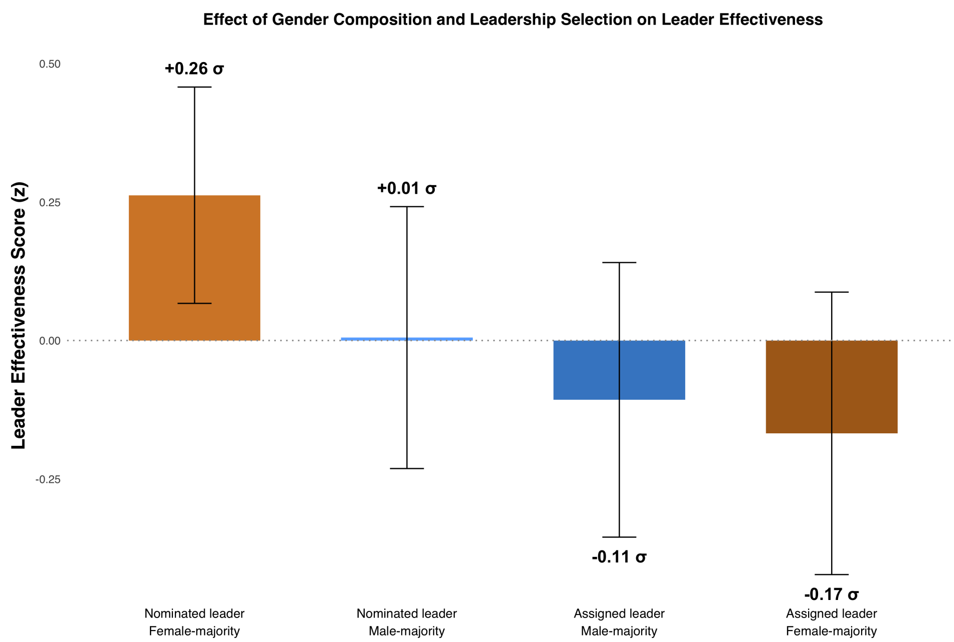


Figure 1.9: *Effect on Gender Composition & Leadership Selection on Leader Effectiveness Rating*

Notes: *This figure shows leader effectiveness scores for each treatment group and pairwise comparisons. Leader effectiveness score is based on a single question, where respondents rate their team leader on a scale of 1-10. The score is calculated from the subset of individuals who are not team leaders, and standardized to have a mean of zero and a standard deviation of one. Each bar shows the mean of leader effectiveness z-score for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

Each figure plots one of the following three measures of leadership: (i) a standardized team leader effectiveness score (1–10 score measured at the endline); (ii) a standardized leadership index composed of five key leadership dimensions— coordination, motivation, conflict resolution, task distribution, and openness to team members' ideas; [8] (iii) teammates' perception of how much (in percentage) of the work do team leaders contribute to. Results are shown in Figure 1.9

[8] I construct a continuous index of perceived leadership effectiveness by extracting standardized factor scores from a one-factor exploratory factor analysis estimated using maximum likelihood. Internal consistency is high, with Cronbach's alpha = 0.93.

(Effectiveness), Figure 1.10 (Leadership Index), and Figure 1.11 (Leader Contribution) with corresponding point estimates and standard errors in tabular form in Table 1.5.

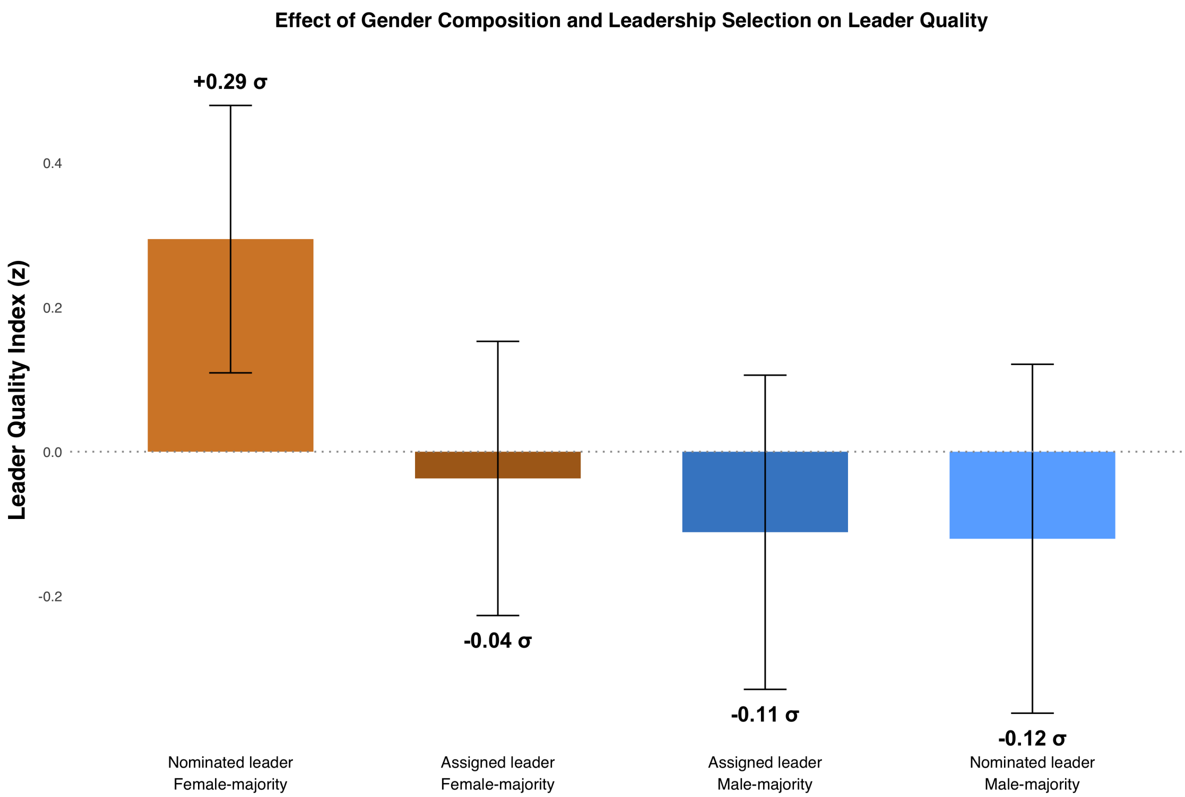


Figure 1.10: *Effect on Gender Composition & Leadership Selection on Leader Quality*

Notes: *This figure shows scores from an index of leadership for each treatment group. Leadership index is constructed using exploratory factor analysis, using five survey items that correspond to different aspects of leadership. The score is calculated from the subset of individuals who are not team leaders and standardized to have a mean of zero and a standard deviation of one. Each bar shows the mean of leadership index z-score for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

In Figure 1.9 and Column 1 of Table 1.5 female-majority teams with nominated leaders are perceived as having the most effective leadership, scoring 0.26σ (p-value < 0.1) higher than male-majority teams with nominated leaders and substantially higher (0.43σ, p-value < 0.01) than female-majority teams with assigned leaders. There are no statistically significant difference in leader effectiveness rating between male and female-majority teams with assigned leaders and between male-majority teams.

Figure 1.10 and Column 2 of Table 1.5 present differences in the leadership index–a holistic measure of perceived leadership quality. Here, too, nominated leaders of female-majority teams are perceived to perform the best. They score 0.42σ (p-value < 0.01) higher on the leadership index than male-majority teams with nominated leaders and 0.33σ (p-value < 0.05) higher than female majority teams with assigned leaders. As before, there are no statistically significant differences between the other groups.

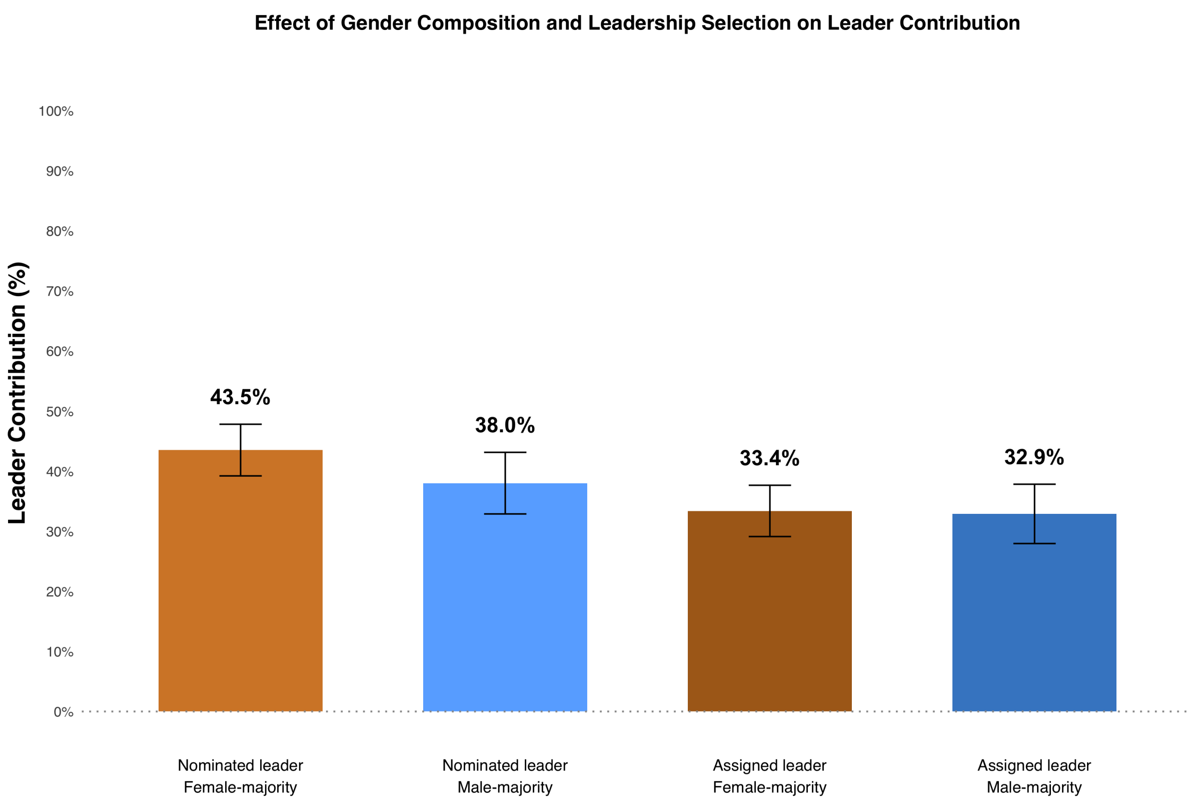


Figure 1.11: *Effect on Gender Composition & Leadership Selection on Leader Contribution*

Notes: *This figure shows leader contribution (in %) for each treatment group. Leader contribution is calculated only from the subset of individuals who are not team leaders, with each team member reporting how much their leader contributed to the team's work. Each bar shows the mean of leader contribution for the indicated group. Error bars represent 95% confidence intervals based on robust standard errors clustered at the team level.*

Finally, peer-reported leader contributions are reported in Figure 1.11 and Column 3 of Table 1.5. As before, the leaders of female-majority teams with nominated leaders are perceived to be the most effective according to this measure. Teams with assigned leaders–male and female-majority–both report their leaders contribute approximately 33% to their collaborative work. Put differently, members of these teams perceive that their leaders contribute their "fair share" (for a 3 member team, equal or a fair share of work would be 33.33%). Teams with nominated leaders do more than their fair share–especially nominated leaders of female-majority teams are perceived to contribute the most on average (43.5%) and significantly more by 8.34pp (p-value < 0.01) than assigned leaders of female-majority teams.

The results across leadership outcomes–leader effectiveness, leader quality, and leader contribution–highlight that leadership selection matters consistently for leaders of female-majority teams. Assigning a leader (vs. nominating one) reduces leader effectiveness by 0.43σ ($p < 0.01$), reduces leader quality by 0.33σ ($p < 0.05$), and decreases peer-reported leader contribution by 8.34 pp ($p < 0.01$) among teams with female-majority. These large and significant differences in leadership between female-majority teams emphasize the importance of leadership selection for female-majority teams. By contrast, among male-majority teams, assignment yields small and statistically insignificant differences in effectiveness (-0.11σ) and quality ($+0.01\sigma$), and a marginally significant difference in leader contribution (-5.36 pp, p-value < 0.1). This seems to suggest that leadership selection is less important for male-majority teams.

Taken together, the leadership results explain the variation in performance between female-majority teams and reinforce the variation in performance between female and male-majority teams with nominated leaders. The result also complement the teamwork outcomes in the earlier subsection: gender composition shapes the general teamwork climate (effectiveness and equity), whereas leadership selection—especially peer nomination—shapes how leaders are evaluated and how much they contribute, with these effects concentrated in female-majority teams. This pattern helps to reconcile why female-majority teams with nominated leaders perform consistently better than other teams. Lastly, the results emphasize the importance of leaders for a team's performance. Specifically, leadership quality (Figure 1.10) – a holistic measure of a leader's perceived performance–closely maps onto the overall performance of teams (Figure 1.5), much more than teamwork-related outcomes. This suggests that while teamwork is important for how individuals perceive the experience of working in a team, leadership seems to correspond more strongly to a team's performance.

1.4.4 *Variation in Characteristics of Team Leaders*

An important question that arises from the differences in leadership' effectiveness, quality, and contributions is whether leaders themselves differ systematically between teams. Specifically, do certain types of leaders-such as those who are nominated versus assigned, or those leading female-majority versus male-majority teams-possess different baseline characteristics that might explain variation in their effectiveness and performance? To examine this question, I report differences in the' background of leaders (gender, caste, and income) and their ability (IQ, emotional intelligence, and preference to lead) across teams. I present these differences in Table 1.6.

Panel A of Table 1.6 compares female- and male-majority teams within each leadership structure. Leaders in female-majority teams, regardless of their leadership structure, are substantially more likely to be women (+46.9 pp & +42.4 pp, $p < 0.001$), consistent with the higher proportion of women in those teams. Beyond gender, there are few systematic differences: leaders in female-majority teams do not differ in socioeconomic background, cognitive ability, or emotional intelligence, though they exhibit a slightly higher preference for leadership (+0.17σ & +0.28σ, for female-majority teams with nominated and assigned leaders, respectively, compared to their male-majority counterparts–though neither are statistically significant).

Table 1.6: *Team Leader Characteristics by Leadership Selection & Gender Composition Group*

	Female (=1)	Upper (=1)	Low income (=1)	Raven (z)	PAGE (z)	Leader pref. (z)
Panel A: Within Leadership Structure						
Nominated leader: Female-majority (=1)	0.469***	0.065	-0.073	0.072	0.218	0.174
	(0.088)	(0.095)	(0.093)	(0.200)	(0.203)	(0.172)
Assigned leader: Female-majority (=1)	0.424***	-0.020	0.153	-0.222	-0.083	0.275
	(0.091)	(0.091)	(0.097)	(0.207)	(0.151)	(0.199)
Panel B: Within Gender Composition						
Female-majority teams: Assigned (=1)	-0.032	-0.100	0.044	-0.268	0.323*	-0.183
	(0.086)	(0.091)	(0.092)	(0.196)	(0.176)	(0.172)
Male-majority teams: Assigned (=1)	0.012	-0.015	-0.182*	0.026	0.625***	-0.284
	(0.094)	(0.095)	(0.098)	(0.210)	(0.181)	(0.199)
Panel C: Full sample						
Female-majority (=1)	0.447***	0.022	0.039	-0.076	0.076	0.220*
	(0.063)	(0.066)	(0.068)	(0.144)	(0.131)	(0.132)
Assigned (=1)	-0.005	-0.059	-0.063	-0.130	0.466***	-0.227*
	(0.070)	(0.066)	(0.067)	(0.143)	(0.126)	(0.131)
Observations (leaders)	203	203	203	203	203	203

Notes: Each cell reports the coefficient from a separate OLS regression of the listed characteristic on the comparison indicator (Panel A: *female-majority* vs *male-majority* within nominated/assigned teams; Panel B: *assigned* vs *nominated* within female-/male-majority teams; Panel C: same two comparisons in the full team-leader sample). Standard errors clustered by team are in parentheses. Intercepts omitted. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Panel B compares leadership structures within each gender composition. Among male-majority teams, assigning a leader—as rather than nominating one—reduces the share of low-income leaders (–18.2 pp, $p < 0.10$) and selects leaders with substantially higher emotional intelligence (+0.63σ, $p < 0.001$). Among female-majority teams, the only difference is in PAGE scores (+0.32σ higher for assigned leaders). The higher PAGE scores among the assigned leaders are partly mechanical: in these arms, the leaders were deliberately chosen based on their emotional intelligence (proxied by PAGE) measured at baseline. Although assigned leaders exhibit lower preference for leadership (-0.18σ & -0.28σ, for female-majority and male-majority teams with assigned leaders, respectively) neither are statistically significant.

These results suggest that while team composition largely shapes who becomes a leader (e.g. increasing representation of women), the leader selection method shapes what kind of leader is chosen (e.g. one with a higher preference to lead or higher emotional intelligence). This distinction may help explain earlier results: nominated leaders, especially in female-majority teams, who are more motivated to lead, tend to contribute more, and are evaluated more favorably by their teammates. Assigned leaders, while more emotionally skilled, may be less inclined to take ownership of the leadership role, which could help explain their lower contributions and effectiveness (observed in earlier tables). This misalignment is also reflected in the low correlation ($\rho = 0.08$) between willingness to lead and emotional intelligence as measured by PAGE.

Interestingly, there are no differences in the gender and caste identity of leaders regardless of whether they are assigned externally or nominated by peers. This is surprising in India's context, where gender and caste play a prominent role in shaping social dynamics. It would appear that other factors, such as the willingness to lead, take precedence in an environment where college teams are required to work collaboratively on projects with strong financial incentives.

1.4.5 *Free Riding*

The analysis so far indicates that female-majority teams benefit from agency in choosing their own leaders, but male-majority teams do not. To better understand this phenomenon, I further examine another source of variation. Specifically, I analyze whether the patterns of free-riding differ by team gender composition and leadership structure. I construct an indicator for *free* *rider* that is equal to 1 if both peers of a team member report that the individual contributed nothing (or 0%) to the project and 0 otherwise. This measure therefore captures cases of complete non-participation as perceived unanimously by teammates. The results of free-riding are presented in Table 1.7.

Table 1.7: *Free-riding by Team Gender Composition and Leader Assignment*

	Free rider (=1)
Panel A: Within Leadership Structure	
Nominated leader: Female-majority (=1)	-0.016
	(0.027)
Assigned leader: Female-majority (=1)	0.030
	(0.019)
Panel B: Within Gender Composition	
Female-majority teams: Assigned (=1)	-0.018
	(0.024)
Male-majority teams: Assigned (=1)	-0.064***
	(0.023)
Observations (individuals)	569

Notes: Each cell reports the coefficient from a separate OLS regression. Heteroskedasticity-robust standard errors clustered at the team level are in parentheses. Dependent variable is an indicator for free-riding (0/1). Intercepts omitted. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Panel A compares female- and male-majority teams within each leadership structure. The estimates show no statistically significant differences between female- and male-majority teams when leaders are nominated: female-majority teams report 1.6 pp fewer free-riders, but not significantly ($p > 0.10$). However, descriptive evidence helps to clarify the pattern. Among teams with nominated leaders, men and women in female-majority teams free-ride at roughly equal rates, whereas in male-majority teams all free-riders are men. This aligns with the idea that in male-majority settings, men who are not chosen as leaders may be less willing to cooperate or contribute. When leaders are externally assigned, the difference in free-riding between female- and male-majority teams remains small and insignificant.

Panel B compares team with the same gender composition. Among female-majority teams, assigning a leader has little effect on free-riding. In contrast, among male-majority teams, assigning a leader based on emotional intelligence significantly reduces free-riding by 6.4 pp ($p < 0.01$). This suggests that in male-majority teams, externally assigning a leader may mitigate within-team conflict or disengagement that arises when peers compete for leadership.

These findings support the mechanism proposed earlier: in male-majority teams that choose their own leaders, male teammates who are not selected appear more likely to withdraw effort, consistent with prior evidence that male-to-male competition can undermine cooperation (Sutter et al., 2009; Eckel & Grossman, 2001; Castillo et al., 2010). In contrast, teams with female-majority show more balanced participation and are less affected by the leadership selection process.

1.4.6 *Contribution and Evaluation of Women Leaders*

Prior literature at the intersection of gender and leadership suggests that a team's gender composition and how leaders are selected shape the representation and role of women leaders in important ways. Certain settings in particular–such as male-dominated fields–are detrimental to women's leadership. In previous sections, I reported a somewhat surprising finding: irrespective of a team's gender composition and leadership selection process, women leaders are well-represented and are not marginalized from leadership roles. However, contingent on women assuming leadership roles, are there differences between how women and men leaders contribute and are evaluated? The answers to this question are presented in Table 1.8.

Table 1.8: *Differences in Perceived Performance of Male & Female Leaders*

	All teams	Female-majority teams	Male-majority teams
Contribution (peer-reported)	5.608**	-0.373	11.710***
	(2.228)	(3.407)	(3.275)
TL effectiveness (peer, z)	0.038	-0.214	0.223
	(0.113)	(0.198)	(0.160)
TL quality index (peer)	0.088	-0.259*	0.201
	(0.103)	(0.145)	(0.172)
Observations (individuals)	371	199	172

Notes: Entries are coefficients on *Female Leader* (vs male leader) from separate OLS regressions. Standard errors clustered by team are in parentheses. Columns split the sample by team gender composition. Intercepts omitted. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

I find that on average, women leaders are perceived to contribute more than male leaders (+5.6 pp, p-value $<$ 0.05). This average masks strong heterogeneity by team gender composition: in male-majority teams, the contribution gap is large and precisely estimated (+11.7 pp, p-value $<$ 0.01), whereas in female-majority teams it is near zero and imprecise. In female-majority teams, female leaders contribute about the same as male leaders but receive slightly lower peer ratings: the leadership quality index is -0.26σ and marginally significant (p-value $<$ 0.10), and effectiveness is also negative but imprecise. In male-majority teams, the pattern changes in contribution—female leaders contribute more (+11.71 pp, p-value < 0.01), but their ratings do not rise accordingly. Although their leader effectiveness (+0.22σ) and leader quality (+0.20σ) ratings are positive, they are statistically indistinguishable from zero. In other words, female leaders appear to get no ratings premium when they contribute substantially more in male-majority teams, and they may face a modest ratings penalty in female-majority teams even when their contribution is comparable.

In Appendix Table 1.A3, I also report within-team rating gaps by rater gender using team fixed effects (female rater minus male rater). Essentially, within the same team, do male and female teammates rate their leaders differently based on the gender of the team leader? I find that across both male-led and female-led teams, these gaps in ratings are small and statistically indistinguishable from zero. In other words, I find no evidence of same-sex favoritism or cross-sex penalties in peer scoring. Taken together with the main table, this suggests that the patterns I document–female leaders contributing more overall (especially in male-majority teams) but not receiving higher leadership ratings, and a modest ratings shortfall for female leaders in female-majority teams despite similar contribution–are not driven by rater-gender bias. Rather, these patterns might indicate that peers, regardless of their gender, might have higher expectations for how much female leaders should contribute to be considered as effective as male leaders. Alternatively, the divergence might also reflect differences in how teams translate observed contributions of leaders into evaluations of leaderships across contexts.

1.4.7 *Robustness Checks*

To assess the robustness of the main results, I examine whether the findings are sensitive to sample restrictions and alternative model specifications. The main specification, presented in Table 1.3, includes year-fixed effects to account for the randomization of teams within academic cohorts.

First, I re-estimate the main outcomes after excluding two teams whose gender composition does not clearly align with the predominant male- or female-majority categories used in the design (the two teams are gender-balanced but coded as female-majority in the main analysis). Table 1.A4 shows that the results remain nearly identical to the baseline specification: female-majority teams with nominated leaders continue to have the highest submission rates and overall performance. The estimated effects are consistent in magnitude and significance in all four outcomes, indicating that the main findings are not driven by these atypical teams.

Next, I assess whether the results are robust to omitting year fixed effects. Although randomization occurred within each academic cohort, removing year fixed effects allows me to check that the results are not solely an artifact of controlling for between-year variation. Table 1.A5 reports the estimates from this specification. The results remain substantively unchanged: female-majority teams with nominated leaders continue to outperform all other groups across all outcomes. They are about 20 percentage points more likely to submit their projects ($p < 0.01$) and achieve total project scores roughly 0.51σ higher than male-majority teams with nominated leaders ($p < 0.01$). Among female-majority teams, those with assigned leaders continue to perform significantly worse (-0.38σ, $p < 0.05$). Across all four outcomes, including the share of teams that reached the finalist stage, the estimates are nearly identical in magnitude and direction to the main specification, confirming that the findings are not dependent on the inclusion of year fixed effects.

Finally, I re-estimate the main specification including additional individual-level controls for socioeconomic status and ability. Specifically, I control for indicators of caste group, household income, leadership preferences, cognitive ability, and emotional intelligence. These variables account for background characteristics that could potentially influence student and team performance. Table 1.A6 presents the results. The inclusion of these controls has little effect on the estimates: female-majority teams with nominated leaders continue to show significantly higher completion rates and overall performance–approximately 0.48σ higher than male-majority teams with nominated leaders ($p < 0.01$)–and female-majority teams with assigned leaders continue to perform worse (-0.36σ, $p < 0.05$). The coefficients remain similar in size and significance across all outcomes, including the probability of reaching the finalist stage.

Overall, the results across these alternative samples and specifications demonstrate that the main findings are highly robust. The advantage of female-majority teams with nominated leaders persists across all models and is not driven by cohort-level differences, outliers, or differences in observable student characteristics.

1.4.8 *Attrition*

There is no attrition for any of the main performance outcomes, since these are collected at the team-level. As such, I have outcome data for all the teams that were part of the study. In addition to team-level data, I also collected individual-level data via the endline survey–these data were then aggregated at the team-level, specifically for outcomes related to teamwork and leadership. Although the overall response rate for the endline is high, (93.3%), I check for differential attrition between treatment groups with 'missingness' as the outcome. I report these results in Table 1.A7. I find differential attrition between teams with nominated leaders. Specifically, male majority teams that nominated their own leader have 7.1pp higher attrition than female-majority teams with a nominated leader. I find no evidence of differential attrition between other treatment groups.

Using Lee (2009) bounds, I re-estimate the differences in teamwork and leadership outcomes between female- and male-majority teams with nominated leaders to assess whether the main results are sensitive to differential attrition. The bounded estimates, reported in Table 1.A8, are very similar to the original coefficients, suggesting that the main findings for teamwork and leadership are robust to correcting for potential selection due to attrition.

1.5 Discussion and Conclusion

This field experiment in India examines the effect of a team's gender composition and method of leadership selection on performance, teamwork, and the representation and effectiveness of leaders. I show that when female-majority teams have the agency to choose their own leaders, they outperform male-majority teams (regardless of leadership structure) and also outperform female-majority teams with leaders assigned based on emotional intelligence. In general, female-majority teams demonstrate better engagement and teamwork, which might reflect more prosocial behavior when women collaborate. The study also shows that the process of leadership selection in mixed-gender teams matter, especially when women are in a majority. The choice of leadership selection impacts the perceived effectiveness, quality, and contribution of leaders, and closely mirrors improvements in a team's performance.

With regards to the role of women in teams, some results of this study complement findings in other settings–more cooperation among women, lower evaluation of women leaders, and the legitimizing effect of democratic selection for women leaders. However, other findings depart from earlier studies in subtle but important ways. For example, I find no evidence of the under-representation of women in leadership roles, regardless of whether they're nominated by peers or assigned based on emotional intelligence, or whether they're in male or female-majority teams. On the surface, this might appear somewhat puzzling. Even in more gender-equal Western contexts, women leaders are under-represented in male-dominated STEM fields. How is it that in an arguably less gender-equal context of India, women are able to assume leadership roles?

One reason could be women's higher preference for leadership compared to men. Expressing this preference during team deliberations might help them assume leadership roles. Additional evidence from the baseline suggests that women assume leadership roles despite gender bias. For example, men in the study report higher levels of hostile sexism compared to women[9]. In one of the questions related to women in the workplace, slightly more than 50% men agree or strongly agree with the statement *Many women ask for special treatment at work and call it "equality"* compared to 21% women. This aligns with the type of prejudices men might have about women in gender-segregated work environments. However, the election of women leaders, especially in male-majority teams, might suggest that women are able to successfully negotiate these prejudices and assume leadership roles.

There could also be differences in the broader context of developing countries that might shape women's preferences. For instance, women might not take their education–and, more generally, their freedom–for granted. Evidence from developing countries suggests that women's educational decisions involve inter-generational intrahousehold bargaining–specifically, in patriarchal cultures that value obedience, women are expected to negotiate with their parents to pursue higher levels of education (Ashraf et al., 2020). Early qualitative work from my field site indicate patterns of intrahousehold bargaining. Women at the college encounter opposition from their families to entering the labor market after graduation, especially since the labor market in technology and engineering would most likely involve moving away from home, living independently, and being

[9]Adapted from the hostile sexism inventory in Glick and Fiske (1996). The inventory consists of a 3-item index, which asked participants to respond on a 5-point Likert scale to questions related to whether women get upset too easily; whether women ask for special treatment at work; and whether women make their work problems seem bigger than they are. Men report 0.46σ (p-value < 0.001) higher hostile sexism than women.

exposed to modern mores. Most women in this context prepare themselves to negotiate with their parents to 'allow' them to enter the labor market. These dynamics for college-going women in India are different from their counterparts in Western countries, and could shape women's motivation and their willing to participate and lead during college in unexpected ways.

Finally, as with any experimental study, external validity is a natural concern. I highlight certain features of the study and the sample to help alleviate those concerns. First, the study's gender mix (50.8% women, 49.2% men) is close to the composition of Indian higher education: women constitute roughly 48% of total higher-education enrollment and about 43% of STEM enrollment in 2021–2022, with national trends showing a gradual rise in women's participation.[10] Second, the institutional context is typical: roughly 60% of colleges in India are located in rural areas, which my field site reflects. Third, the team task mirrors the environments many college graduates enter—mixed-gender, project-based teams—so the core study design (who teams are composed of and how leaders are selected) maps directly onto early-career work settings. While generalization beyond short-horizon student teams—especially to higher-stakes professional settings—should be made with caution, the study provides novel evidence from a developing country context of how team gender composition and leadership selection interact to shape collaboration and performance.

[10] Based on the 'All India Survey on Higher Education', AISHE 2021-2022. While more recent data are unavailable, trends point to a gradual increase in the enrollment of women.

Tables

Table 1.4: *Teamwork by Gender Composition and Leadership Structure*

	Team effectiveness	Team equity
Panel A: Within Leadership Structure		
Nominated leader: Female-majority (=1)	0.336**	0.318**
	(0.134)	(0.125)
Assigned leader: Female-majority (=1)	0.178	0.209*
	(0.138)	(0.127)
Panel B: Within Gender Composition		
Female-majority teams: Assigned (=1)	-0.144	-0.121
	(0.127)	(0.103)
Male-majority teams: Assigned (=1)	0.014	-0.012
	(0.137)	(0.141)
Observations (individuals)	569	569

Notes: Each cell reports the coefficient on the listed indicator from a separate OLS regression with year fixed effects. "Team Effectiveness" and "Team Equity" are both indices constructed from five questions each in the endline survey, standardized to have a mean of 0 and standard deviation of 1. Heteroskedasticity-robust standard errors, clustered at the team level, are in parentheses. Intercepts omitted. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.5: *Leadership by Gender Composition and Leadership Structure*

	Leader effectiveness	Leader quality	Leader contribution (%)
Panel A: Within Leadership Structure			
Nominated leader: Female-majority (=1)	0.257*	0.415***	4.359
	(0.144)	(0.139)	(2.952)
Assigned leader: Female-majority (=1)	-0.060	0.075	1.379
	(0.173)	(0.150)	(3.347)
Panel B: Within Gender Composition			
Female-majority teams: Assigned (=1)	-0.430***	-0.332**	-8.349***
	(0.160)	(0.131)	(2.986)
Male-majority teams: Assigned (=1)	-0.112	0.009	-5.369*
	(0.150)	(0.153)	(3.057)
Observations	371	371	371

Notes: This only includes the subsample of individuals who are not team leaders. Peers–who are not team leaders–evaluate team leaders on these three outcomes. Each cell reports the coefficient on the listed indicator from a separate OLS regression with year fixed effects. Heteroskedasticity-robust standard errors, clustered at the team level, are in parentheses. Intercepts omitted. "Leader Effectiveness" is based on a survey question that asked respondents to rate their team leader on a scale of 1-10, which here is standardized to have a mean of 0 and a standard deviation of 1. "Leader Quality" is an index based on five survey questions that capture various components of leadership such as coordination, motivation, conflict resolution, task distribution, and openness to team members' ideas. It is standardized to have a mean of 0 and a standard deviation of 1. "Leader contribution" is expressed as a percentage of total team contributions. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Chapter 1 Appendix

1.A Additional Tables and Figures

Table 1.A1: *Covariate Balance by Gender and Team Composition*

	Female				Male			
Variable	Maj. Female	Maj. Male	Diff	p-value	Maj. Female	Maj. Male	Diff	p-value
Hindu	0.958	0.979	-0.021	0.317	0.917	0.932	-0.015	0.698
Low income	0.701	0.792	-0.091	0.203	0.574	0.620	-0.046	0.268
Mother completed school	0.243	0.188	0.055	0.249	0.343	0.292	0.051	0.299
PAGE score	14.893	14.667	0.226	0.320	14.500	14.708	-0.208	0.195
Raven score	5.332	5.146	0.186	0.541	5.500	5.740	-0.240	0.257
Upper caste	0.272	0.263	0.009	0.981	0.333	0.279	0.054	0.403
Assigned leader	0.500	0.490	0.010	0.766	0.509	0.490	0.020	0.704
Preference for leading	7.173	6.938	0.235	0.399	6.222	5.911	0.311	0.543
Observations	214	96			108	192		

Notes: The table reports mean values of baseline covariates by student gender and team gender composition. Differences are from t-tests with year fixed effects. Stars to indicate statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 1.A2: *Covariate Balance by Team Leader Selection Groups*

Variable	**Mean (Nominated)**	**Mean (Assigned)**	**Difference**	**p-value**
Female	0.507	0.508	0.001	0.969
Female-majority team	0.520	0.535	0.015	0.831
Hindu	0.955	0.936	−0.019	0.304
Low income	0.678	0.657	−0.021	0.587
Mother completed school	0.274	0.261	−0.013	0.716
PAGE score	14.804	14.673	−0.131	0.697
Raven score (std.)	0.109	−0.110	−0.219	0.119
Team lead pref. (std.)	0.022	−0.022	−0.045	0.751
Upper caste	0.287	0.281	−0.006	0.866
Year 1	0.618	0.634	0.016	0.815
Year 2	0.275	0.267	−0.007	0.909
Year 3	0.108	0.099	−0.009	0.837
Observations	*102*	*101*		

Notes: The table reports group means and t-test p-values comparing teams with assigned versus nominated leaders. Stars denote significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Joint F-test p-value for all covariates: 0.541.

Table 1.A3: *Within-team rating gap (Female rater – Male rater) by leader sex*

	Male leader	Female leader	Diff-in-gaps (F–M)
TL effectiveness (peer, z)	0.018 (0.278)	0.227 (0.197)	0.210 (0.341)
TL quality index (peer)	0.293 (0.280)	0.061 (0.221)	-0.231 (0.357)

Notes: Entries are within-team rating gaps (Female rater – Male rater) from OLS with team fixed effects; standard errors clustered by team in parentheses. The "Female leader" column equals the male-leader gap plus the FemaleLeader×FemaleRater interaction. "Diff-in-gaps (F–M)" is the interaction coefficient. Intercepts omitted.

Table 1.A4: *Performance by Gender Composition and Leadership Structure (Drop tie teams)*

	Submitted	Project score (z)	Project score (z), 0 for non-submit	Finalists (%)
Panel A: Within Leadership Structure				
Nominated leader: Female-majority (=1)	0.211*** (0.080)	0.170 (0.221)	0.510*** (0.190)	0.123** (0.056)
Assigned leader: Female-majority (=1)	-0.014 (0.090)	0.279 (0.262)	0.098 (0.213)	0.039 (0.066)
Panel B: Within Gender Composition				
Female-majority teams: Assigned (=1)	-0.174** (0.074)	-0.041 (0.182)	-0.374** (0.176)	-0.020 (0.063)
Male-majority teams: Assigned (=1)	0.050 (0.092)	-0.150 (0.278)	0.038 (0.216)	0.064 (0.053)
Observations (teams)	201	155	201	201

Notes: Each cell reports the coefficient on the listed indicator from a separate OLS regression with year fixed effects. Heteroskedasticity-robust standard errors clustered at the team level are in parentheses. Intercepts omitted. "Finalists" is an indicator for Top-20 teams. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A5: *Performance by Gender Composition and Leadership Structure (Without YFE)*

	Submitted	Project score (z)	Project score (z), 0 for non-submit	Finalists (%)
Panel A: Within Leadership Structure				
Nominated leader: Female-majority (=1)	0.212***	0.159	0.509***	0.090*
	(0.077)	(0.219)	(0.185)	(0.054)
Assigned leader: Female-majority (=1)	-0.016	0.268	0.087	0.005
	(0.088)	(0.252)	(0.208)	(0.062)
Panel B: Within Gender Composition				
Female-majority teams: Assigned (=1)	-0.178**	-0.028	-0.383**	-0.020
	(0.073)	(0.181)	(0.174)	(0.063)
Male-majority teams: Assigned (=1)	0.051	-0.137	0.039	0.066
	(0.092)	(0.281)	(0.217)	(0.053)
Observations (teams)	203	156	203	203

Notes: Each cell reports the coefficient on the listed indicator from a separate OLS regression without year fixed effects. Heteroskedasticity-robust standard errors clustered at the team level are in parentheses. Intercepts omitted. "Finalists" is an indicator for Top-20 teams. ${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$.

Table 1.A6: *Performance by Gender Composition and Leadership Structure (With Controls)*

	Submitted	Project score (z)	Project score (z), 0 for non-submit	Finalists (%)
Panel A: Within Leadership Structure				
Nominated leader: Female-majority (=1)	0.206***	0.112	0.475**	0.126**
	(0.079)	(0.204)	(0.186)	(0.054)
Assigned leader: Female-majority (=1)	-0.007	0.261	0.100	0.036
	(0.088)	(0.262)	(0.208)	(0.064)
Panel B: Within Gender Composition				
Female-majority teams: Assigned (=1)	-0.165**	-0.039	-0.356**	-0.016
	(0.071)	(0.179)	(0.170)	(0.060)
Male-majority teams: Assigned (=1)	0.048	-0.189	0.019	0.074
	(0.092)	(0.272)	(0.215)	(0.053)
Observations (teams)	203	156	203	203

Notes: Each cell reports the coefficient on the listed indicator from a separate OLS regression that includes controls. Heteroskedasticity-robust standard errors clustered at the team level are in parentheses. Intercepts omitted. "Finalists" is an indicator for Top-20 teams. ${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$.

Table 1.A7: *Attrition by Team Gender Composition and Leadership Structure*

	Missing (=1)
Panel A: Within Leadership Structure	
Nominated leader: Female-majority (=1)	-0.071**
	(0.034)
Assigned leader: Female-majority (=1)	0.011
	(0.027)
Panel B: Within Gender Composition	
Female-majority teams: Assigned (=1)	0.030
	(0.025)
Male-majority teams: Assigned (=1)	-0.052
	(0.036)
Observations	569

Notes: Each cell reports the coefficient from a separate OLS regression. Heteroskedasticity-robust standard errors clustered at the team level are in parentheses. The dependent variable is an indicator for missingness (1 = missing). Intercepts omitted. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A8: *Lee (2009) bounds for nominated-leader sample (female − male)*

Outcome	Lower bound [95% CI]	Upper bound [95% CI]	Sel. rate (F)	Sel. rate (M)	Trimmed group	Trim fraction α
Team effectiveness (index)	0.345 [0.147, 0.552]	0.544 [0.254, 0.772]	0.963	0.891	Female-majority	0.074
Team equity (index)	0.266 [0.065, 0.485]	0.483 [0.292, 0.661]	0.963	0.891	Female-majority	0.074
Leader effectiveness (std)	0.241 [0.057, 0.433]	0.379 [0.157, 0.562]	0.963	0.891	Female-majority	0.074
Leadership quality (index)	0.340 [0.151, 0.525]	0.538 [0.258, 0.743]	0.963	0.891	Female-majority	0.074

Notes: Bounds are for the nominated-leader subsample; contrast is female-majority minus male-majority. 95% CIs are from a cluster bootstrap by team. The selection indicator equals 1 when the outcome is observed (non-missing). The higher-retention arm is trimmed to match the other arm.

Table 1.A9: *Balance (Male-majority teams): Assigned vs Nominated leader*

	Group means		Test	
Variable	Mean: Nominated	Mean: Assigned	Diff (A–N)	p-value
Student & household				
Female (=1)	0.333	0.333	0.000	0.330
Hindu (=1)	0.952	0.940	-0.013	0.627
Low income (=1)	0.728	0.624	-0.104*	0.069
Mother completed school (=1)	0.252	0.262	0.011	0.831
Aptitude & preferences				
Page short total score	14.653	14.745	0.092	0.855
Raven total score (z)	0.163	−0.077	-0.240	0.233
Team leader preference (z)	−0.164	−0.176	-0.012	0.952
Upper caste (=1)	0.269	0.280	0.011	0.830
Cohort (dummies)				
Year 1 (=1)	0.776	0.787	0.012	0.891
Year 2 (=1)	0.204	0.191	-0.013	0.879
Year 3 (=1)	0.020	0.021	0.001	0.977
Observations (teams)	49	47	–	–

Notes: Entries are means in each group. "Diff (A–N)" reports Assigned minus Nominated with significance stars. Two-sided p-values from OLS of each covariate on an Assigned indicator. Joint F-test (all covariates): $p = 0.656$. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A10: *Balance (Female-majority teams): Assigned vs Nominated leader*

	Group means		Test	
Variable	Mean: Nominated	Mean: Assigned	Diff (A–N)	p-value
Student & household				
Female (=1)	0.668	0.667	-0.002	0.324
Hindu (=1)	0.958	0.936	-0.022	0.397
Low income (=1)	0.632	0.673	0.041	0.452
Mother completed school (=1)	0.294	0.269	-0.025	0.628
Aptitude & preferences				
Page short total score	14.943	14.635	-0.309	0.500
Raven total score (z)	0.059	-0.093	-0.152	0.445
Team leader preference (z)	0.194	0.161	-0.033	0.865
Upper caste (=1)	0.303	0.282	-0.021	0.681
Cohort (dummies)				
Year 1 (=1)	0.472	0.500	0.028	0.774
Year 2 (=1)	0.340	0.346	0.007	0.944
Year 3 (=1)	0.189	0.154	-0.035	0.640
Observations (teams)	53	54	–	–

Notes: Entries are means in each group. “Diff (A–N)” reports Assigned minus Nominated with significance stars. Two-sided p-values from OLS of each covariate on an Assigned indicator. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

1.B Photographs of Competition and App Examples

Figure 1.B1: *In-person Competition Featuring the Top-20 Teams*

सच कहने का साहस और सलीका

राज एक्सप्रेस

नवीन बार्चे

जवाहरलाल इंस्टीट्यूट ऑफ टेक्नालॉजी बोरावां में, हावर्ड यूनिवर्सिटी की रिसर्च द्वारा आयोजित...

रूरल इण्डिया एप आईडिया चैलेंज में 203 टीमों के 609 इंजीनियर्स ने लिया हिस्सा

प्रथम स्थान के विजेता के 15 हजार, द्वितीय स्थान 12 हजार एवं तृतीय स्थान प्राप्त करने वाली टीम ने जीता 10 हजार का इनाम

20 टीमों के 60 इंजीनियर्स ने अंतिम चरण की प्रतियोगिता में जीते पुरस्कार

पूर्व केन्द्रीय कृषि राज्यमंत्री अरूण यादव ने विजेताओं को दी शुभकामनाएं

पूर्व केन्द्रीय कृषि राज्यमंत्री ने विजेताओं को दी बधाई

शुरू होने से यात्रियों की परेशानी समाप्त होगी।

जवाहरलाल इंस्टीट्यूट ऑफ टेक्नालॉजी बोरावां में हावर्ड यूनिवर्सिटी

रूरल इण्डिया एप आईडिया चैलेंज में 203 टीमों के 609 इंजीनियर्स ने लिया हिस्सा

नवभारत न्यूज

खरगोन, निप्र। जवाहरलाल इंस्टीट्यूट ऑफ टेक्नालॉजी बोरावां में शनिवार 12 जुलाई को हावर्ड यूनिवर्सिटी की रिसर्च टीम द्वारा इंजीनियरिंग पाठ्यक्रमों के विद्यार्थियों के लिए रूरल इण्डिया एप आईडिया चैलेंज प्रतियोगिता का आयोजन किया गया।

इस प्रतियोगिता की शुरूआत मई 2025 में एक बेसलाईन सर्वे से हुई थी । जिसमें बी.टेक. और एम.सी.ए. पाठ्यक्रमों कुल 609 विद्यार्थियों ने भाग लिया था । तत्पश्चात उन्हें 3-3 के समूह में कुल 203 टीमों में विभाजित किया गया । हावर्ड यूनिवर्सिटी के रिसर्चर आशुतोष बुराड़िया ने जानकारी देते हुए बताया कि हमारी रिसर्च टीम वर्क और लीडरशिप पर केन्द्रित है । इसके साथ ही विद्यार्थियों की तकनीकी क्षमताओं, टीम बिल्डिंग, लीडरशिप, एप्लीकेशन डेवलपमेंट एवं तकनीकी ज्ञान का आंकलन कर सही दिशा में उपयोग पर केन्द्रित करना है । श्री बुराड़िया ने आगे कहा कि रूरल इण्डिया पर आधारित प्रोजेक्ट को प्रेजेन्टेशन के माध्यम से सभी टीमों द्वारा तैयार कर जमा करना था । हमारी टीम द्वारा इनमें से 20 यूनिक प्रोजेक्ट का चयन किया गया जिसका आज शनिवार को जेआयटी बोरावां में आयोजित प्रतियोगिता में ऑनलाईन प्रेजेन्टेशन के माध्यम से चयन किया गया । इस प्रक्रिया में रिसर्च टीम का हिस्सा रही रिसर्च असिस्टेंट सिद्रा अली एवं 2 निर्णायक तेजस एवं मृदुल ने बेस्ट 3 आईडियाज का ऑनलाईन चयन किया गया ।

प्लेसमेंट अधिकारी एवं कार्यक्रम संयोजक प्रो. नितिन नामदेव ने जानकारी देते हुए बताया कि प्रथम पुरस्कार संस्था की बी.टेक.प्रथम वर्ष की टीम उम्मेहानी , कार्तिकेय और अमन को रू 15 हजार नगद प्रदान किये गये । द्वितीय स्थान पर प्रथम वर्ष की टीम प्रियंका नकुल और ऑकित को राशि रू 12 हजार नगद प्रदान किये । तृतीय स्थान पर अक्षा, सुफियान और नैना की टीम को राशि रू 10 हजार नगद प्रदान किये । इसके साथ ही 17 टीमों को सांत्वना पुरस्कार के रूप में नगद 3 हजार प्रत्येक टीम को प्रदान किये । साथ ही हावर्ड यूनिवर्सिटी की रिसर्च टीम द्वारा प्रतिभागिता प्रमाण पत्र भेजे जावेंगें ।

संस्थान के प्राचार्य डॉ. अतुल उपाध्याय , वरिष्ठ प्राध्यापकों एवं सभी विभागाध्यक्षों व शिक्षकों ने विजेता टीमों को पुरस्कार सौंपकर उनके उज्जवल भविष्य की कामनाएँ की ।इस आयोजन में डॉ. संजय चौहान, डॉ. संतोष पँवार, डॉ. मनीष देशमुख , प्रो. कमलेश पाटीदार, प्रो. विनय देडलकर , प्रो. लवकेश पाटीदार का सहयोग सराहनीय रहा।

पूर्व केन्द्रीय कृषि राज्यमंत्री ने विजेताओं को दी बधाई

पूर्व केन्द्रीय कृषि राज्यमंत्री अरूण यादव ने अपने संदेश में संस्था में शनिवार को आयोजित रूरल इण्डिया एप आईडिया चैलेंज प्रतियोगिता के विजेताओं को बधाई एवं शुभकामनाएं देते हुए कहा कि आज का दौर आधुनिक सोच को क्रियान्वित करने का युग है ।हमारे विद्यार्थियों ने एप्लीकेशन टूल्स का उपयोग करते हुए शिक्षा, कृषि, प्रबंधन, स्मार्ट वर्क, डाटाबेस के प्रोजेक्ट के रूप में एप डिजाईन किये गये ।सभी विजेताओं को मैं बधाई एवं शुभमानाएँ देता हूँ ।आने वाला समय आपका है , आप सभी मेहनत व लगन से पढ़ाई कर संस्थान और अपने परिवार को गौरवान्वित करें ।मैं हावर्ड यूनिवर्सिटी के रिसर्चर आशुतोष बुराड़िया एवं उनकी टीम को भी धन्यवाद करना चाहता हूँ कि आपने हमारे विद्यार्थियों को आधुनिक तकनीकी के क्षेत्र में एक नया मार्ग दिखाया है ।

Figure 1.B2: *Local Media Coverage of the Competition*

1.C Survey Instruments

PAGE Assessment

I constructed a measure of emotional perceptiveness called PAGE (Perceiving AI Generated Emotions) with 28 items for India's context. The assessment is based on Weidmann and Xu (2024), who designed the assessment for a sample in the United States. PAGE was constructed using the prompts from Weidmann and Xu (2024) but using Indian faces and expression. For example, I used the following prompt to generate the face in Figure 1.C3: "Create a hyper-realistic image of a young Indian women showing expression anxiety. Eyes looking sideways, frowned eyebrows, biting lips. Detailed skin texture and natural lighting. Wearing a white tshirt. No body language, showing only the face, head oriented at the front, and staring at the camera. Plain grey background." I adapted prompts for the original paper to generate a range expression for Indian faces. I provide examples of the PAGE assessment in this section.

In terms of its psychometric properties, the PAGE has a moderate Cronbach's Alpha of 0.70. The tests shows good psychometric properties in the other respects as well. For instance, the distribution of PAGE total scores approximately normal (M = 14.7, SD = 3.37, skew = -0.01, kurtosis = -0.18). Fewer than 1% of participants scored at either the minimum or maximum value, indicating no floor or ceiling effects. The distribution of PAGE scores are shown in Figure 1.C1. I also conduct a principal component analysis of PAGE. The principal component analysis (PCA) indicated a strong first component (eigenvalue = 2.1), with a clear 'elbow' in the scree plot after the first component, as depicted in Figure 1.C2. Subsequent components each accounted for substantially less variance. This pattern supports a unidimensional structure of the PAGE, consistent with the intended measure of a single underlying construct of emotional intelligence.

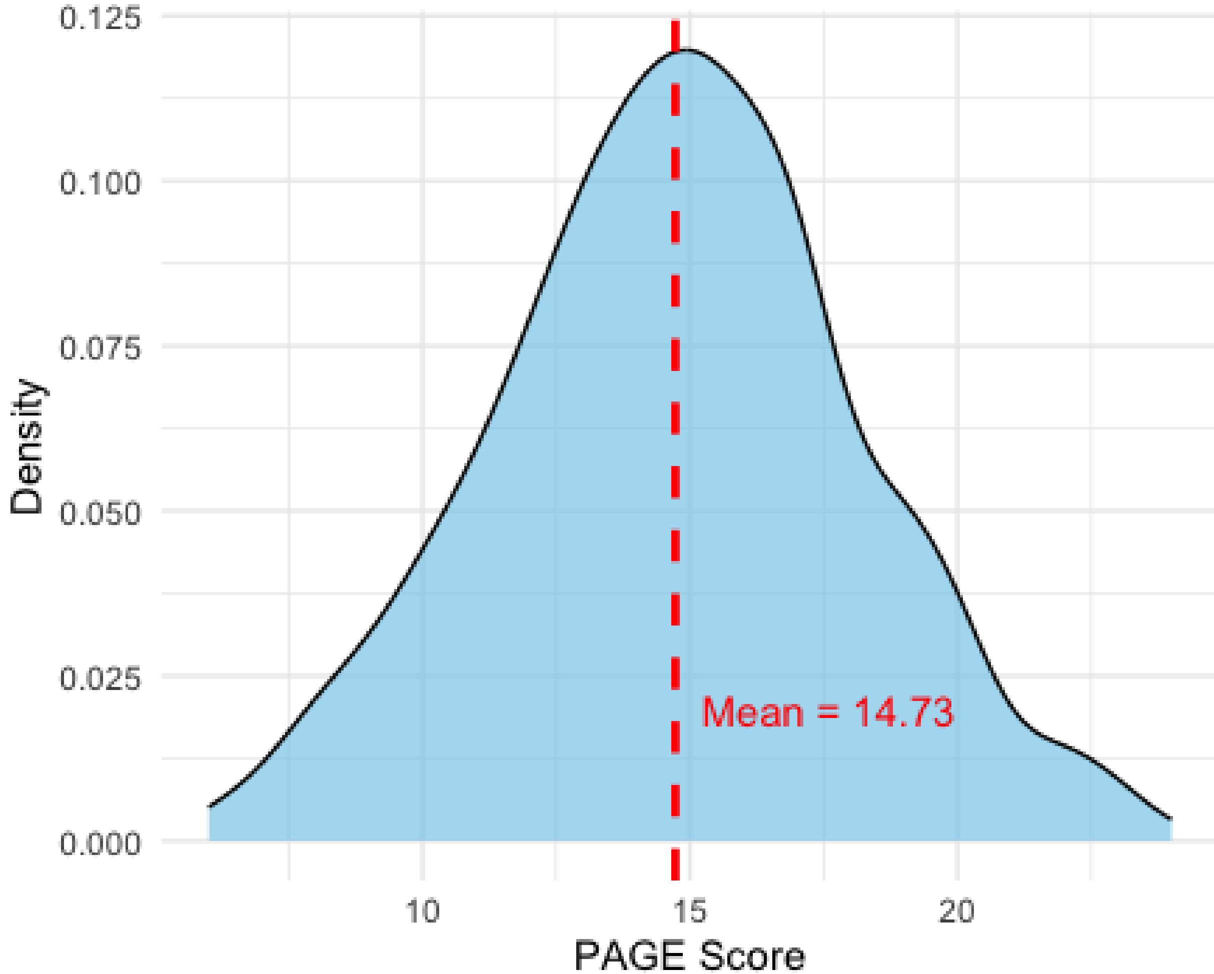


Figure 1.C1: *Distribution of PAGE Score*

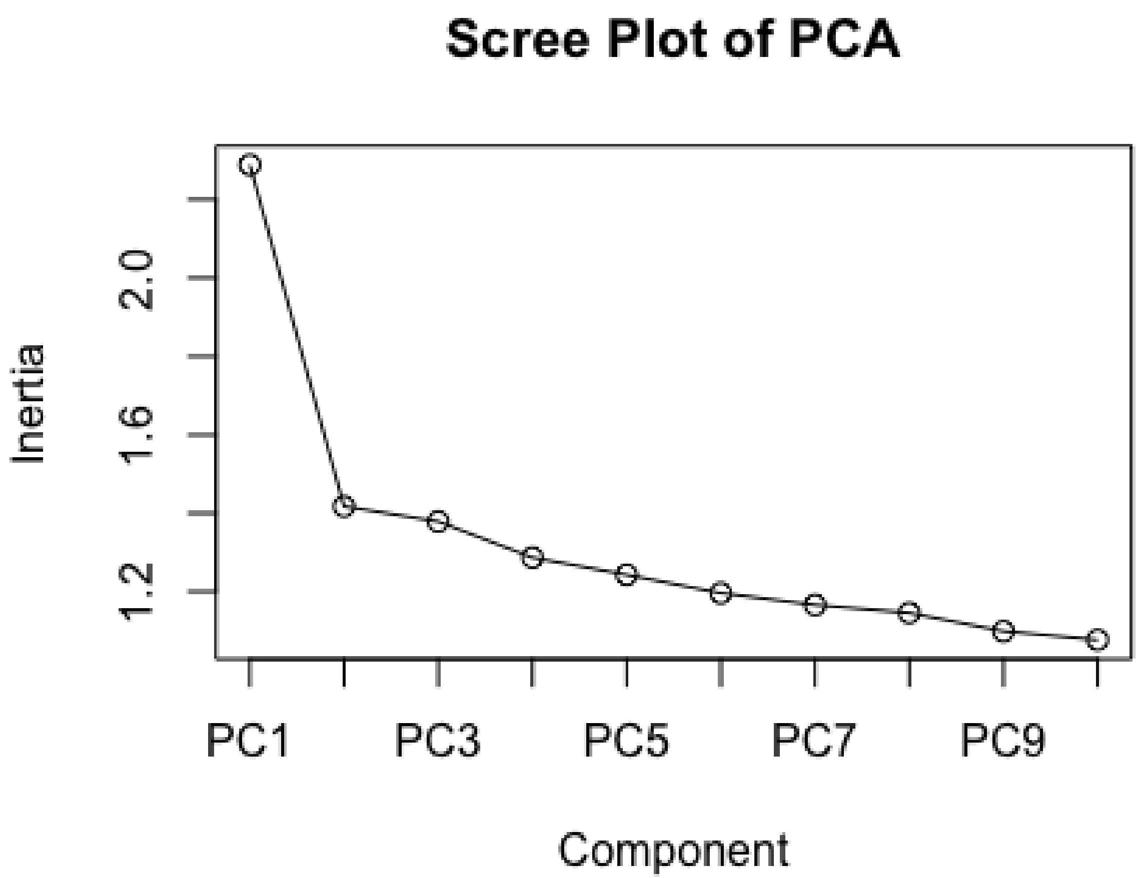
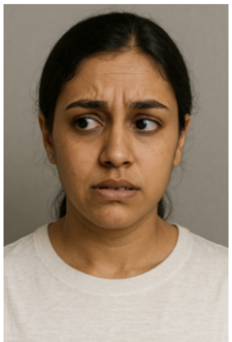


Figure 1.C2: *Scree Plot of PCA for PAGE*

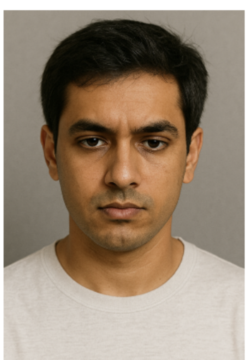
What is the person in the picture thinking or feeling? तस्वीर में दिख रहा व्यक्ति क्या सोच रहा है या क्या महसूस कर रहा है?



- ◯ Embarrassment - शर्मिंदगी
- ◯ Fear - डर
- ◯ Anxiety - चिंता
- ◯ Contemplation - चिंतन - deep thinking
- ◯ Contentment - संतुष्टि - satisfied
- ◯ Confusion - भ्रम - not understanding

Figure 1.C3: *PAGE Item Example 1*

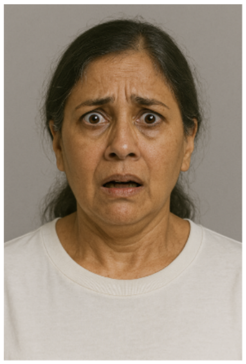
What is the person in the picture thinking or feeling? तस्वीर में दिख रहा व्यक्ति क्या सोच रहा है या क्या महसूस कर रहा है?



- ◯ Disappointment - निराशा
- ◯ Concentration - एकाग्रता - focused
- ◯ Confusion - भ्रम - not understanding
- ◯ Interest - रुचि - wanting to know more
- ◯ Doubt - संदेह - unsure
- ◯ Contentment - संतुष्टि - satisfied

Figure 1.C4: *PAGE Item Example 2*

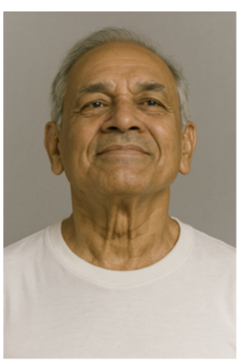
What is the person in the picture thinking or feeling? तस्वीर में दिख रहा व्यक्ति क्या सोच रहा है या क्या महसूस कर रहा है?



- ○ Shame - लज्जा - guilty feeling
- ○ Confusion - भ्रम - not understanding
- ○ Fear - डर
- ○ Anxiety - चिंता
- ○ Awe - विस्मय - amazed
- ○ Surprise - आश्चर्य

Figure 1.C5: *PAGE Item Example 3*

What is the person in the picture thinking or feeling? तस्वीर में दिख रहा व्यक्ति क्या सोच रहा है या क्या महसूस कर रहा है?



- ○ Awe - विस्मय - amazed
- ○ Joy - हर्ष - very happy
- ○ Amusement - मनोरंजन
- ○ Contempt - तिरस्कार
- ○ Pride - गर्व
- ○ Interest - रुचि

Figure 1.C6: *PAGE Item Example 4*

Survey Instruments: Teamwork and Leadership

Team Equity Index. To capture students' perceptions of fairness and inclusion within teams, I constructed a *Team Equity Index* using six Likert-scale items: (i) my ideas were valued by the team; (ii) I was able to contribute meaningfully to the project; (iii) workload was distributed equitably; (iv) all team members had equal opportunity to contribute; (v) members from all backgrounds were equally respected; and (vi) I felt comfortable and included within the team. Responses were coded from 1 (Strongly disagree) to 5 (Strongly agree). The index was derived from a one-factor exploratory factor analysis with regression-based scoring and standardized to mean zero and unit variance. Internal consistency is high ($\alpha = 0.80$).

Team Effectiveness Index. The *Team Effectiveness Index* measures perceived team functioning and performance. It is based on five items: (i) our team worked together effectively; (ii) our team made good use of time and resources; (iii) team members communicated well; (iv) our team resolved conflicts successfully; and (v) I am satisfied with our team's performance. Each item was rated from 1 (Strongly disagree) to 5 (Strongly agree). Scores were extracted using one-factor exploratory factor analysis and standardized (mean = 0, SD = 1). Reliability is high ($\alpha = 0.92$).

Leadership Index. To assess leadership quality, I constructed a *Leadership Index* from five items capturing coordination, motivation, conflict resolution, task distribution, and openness to team members' ideas. Each team member rated the team leader on these dimensions from 1 (Strongly disagree) to 5 (Strongly agree). A single-factor model was estimated using exploratory factor analysis to produce standardized factor scores. The index has strong internal consistency ($\alpha = 0.93$).

Leader Effectiveness Score. Participants were also asked "On a scale of 1-10, how would you rate your team leader's overall effectiveness?" 1 = not effective at all and 10 = extremely effective.

1.D Project Evaluation Rubric

App Project Evaluation Rubric

Total Points: 20 (4 points per criterion)

Problem Definition Clarity (4 points)

4 - Excellent

- Clearly identifies a specific rural problem with context or data/evidence.
- Shows understanding of target users and their basic challenges.
- Problem statement is rural-focused and well-defined.
- Demonstrates awareness that this is an issue worth addressing.

3 - Proficient

- Identifies a legitimate rural problem, though articulation may lack depth.
- Shows basic understanding of target users.
- Problem statement is clear enough to understand what they aim to solve and why it matters.
- Makes connection to the rural context, even if general.

2 - Developing

- Identifies a problem that could be relevant to rural areas, though the statement is general.
- Limited understanding of specific user needs or context.
- Problem statement lacks specificity but the core issue is understandable.
- Weak attempt to connect to the rural setting.

1 - Beginning

- Problem statement is unclear, overly broad, or not rural-specific.
- Little to no understanding of target users or context.
- Fails to articulate why this problem matters.
- No clear connection to rural challenges.

Solution Feasibility, Especially in Rural Context (4 points)

4 - Excellent

- Shows strong understanding of rural constraints (connectivity, literacy, devices).
- Most features are realistic and appropriate for rural implementation.
- Includes adaptations for rural context (simple interface, offline features, etc.).
- Demonstrates awareness of the limitations of an app-based solution.
- Feasible with only minor adjustments needed.

3 - Proficient

- Shows basic awareness of rural constraints.
- Most features can be implemented with basic technology.
- Considers rural constraints in design.
- Demonstrates moderate awareness of limitations.
- Feasible with moderate adaptation.

2 - Developing

- Limited consideration of rural constraints, but not entirely unrealistic.
- Some features would be challenging to implement in rural settings.
- Would need significant adaptation, but core idea has potential.

- Shows superficial awareness of limitations.

1 - Beginning

- Little to no consideration of rural constraints.
- Features are unrealistic for rural deployment.
- Core idea lacks feasibility or potential.
- Shows no awareness of limitations.

UI Simplicity & Accessibility (4 points)

4 - Excellent

- Interface is clean, simple, and user-friendly.
- Effective use of icons, visual cues, voice prompts, and local language elements.
- Clear and logical navigation with minimal cognitive load.
- Considers different literacy levels and accessibility.
- Interface would be usable by rural users.

3 - Proficient

- Interface is functional and intuitive.
- Includes visual elements supporting usability.
- Navigation is logical, though some tasks may require extra steps.
- Interface adequate for target users.

2 - Developing

- Interface functional but somewhat complex or unclear.
- Limited use of visual aids or accessibility features.

- Navigation possible but not intuitive.
- May challenge some users.

1 - Beginning

- Minimal or unclear information on user interaction.
- Heavy reliance on text without visual support.
- Navigation is confusing or illogical.
- No consideration of literacy or accessibility.

Technical Innovation (within constraints) (4 points)

4 - Excellent

- Includes more than two creative or unique features.
- Thoughtful use of technologies (QR code, photo input, voice, SMS, offline mode, etc.).
- Demonstrates original thinking within no-code constraints.
- Shows clear understanding of technical possibilities.

3 - Proficient

- Has one or two creative or unique features.
- Uses available technologies effectively.
- Shows creativity, though somewhat conventional.
- Demonstrates good understanding of technical constraints.

2 - Developing

- Mostly conventional but functional.
- Limited creativity or feature variety.

- Some understanding of technical possibilities.

1 - Beginning

- Basic or conventional with no unique features.
- Minimal use of technology.
- Little evidence of innovation or understanding of constraints.

Potential Social Impact (4 points)

4 - Excellent

- Demonstrates meaningful, measurable potential impact (possibly with numbers).
- Shows understanding of how the solution addresses root causes.
- Considers sustainability, scalability, and long-term effects.
- Includes realistic mechanisms for adoption and behavior change.

3 - Proficient

- Explains potential benefits to users clearly.
- Shows the solution addresses a real need.
- Sets reasonable expectations for impact.
- Includes a basic plan for user adoption.

2 - Developing

- Describes potential benefits vaguely.
- Shows limited understanding of how impact would occur.
- Weak connection between problem and benefits.
- Minimal consideration of sustainability or adoption.

1 - Beginning

- Benefits are unclear or unrealistic.
- Fails to explain how positive change would occur.
- Poor connection between problem and proposed solution.
- No mention of adoption or implementation.

Chapter 2

Misperceptions about Caste and Attitudes toward Affirmative Action Among Youth in India[1]

2.1 Introduction

Affirmative action is a widely used redistributive policy to enhance equity and representation in higher education. There is some form of affirmative action in higher education admissions in a quarter of the world's countries (Post, 2016). Affirmative action can be on the basis of race, ethnicity, gender, class, and caste. In India, affirmative action in universities (termed "reservations") takes the form of quotas for students from lower-caste groups. More than half of the admission spots in public universities can be "reserved" for students from lower-caste groups (Sönmez and Yenmez, 2019). Given that affirmative action policies potentially redistribute educational resources from upper-caste students to their lower-caste counterparts, these policies are highly contentious in India (Deshpande, 2016).

Upper-caste youth in India—who are not the beneficiaries of affirmative action policies—might have varying degrees of opposition to the policy of affirmative action. In part, their opposition might be based on beliefs about the social and economic status of lower-caste groups. If upper-caste youth misperceive lower-caste groups' social and economic status, they might be less inclined to support equity-enhancing policies such as affirmative action. For instance, if upper-caste youth underestimate income disparities between caste groups and underestimate caste-based discrimination, they might be less supportive—and even resentful—of policies that give lower-caste students a "leg up".

Similarly, lower-caste youth might also have varying misperceptions about caste that shape

[1]I thank Emmerich Davies, Jennifer Hochschild, Peter Blair, Saloni Gupta, Alex Bolves, Nishith Prakash, and Stefanie Stantcheva for their helpful comments and suggestions. This research was supported by a Stone Research Grant from Harvard Kennedy School's James M. and Cathleen D. Stone Program in Wealth Distribution, Inequality, and Social Policy. Harvard IRB approval IRB23-0502. This study is registered in the AEA RCT Registry (AEARCTR-0011706).

their preferences for affirmative action and attitudes toward caste. For instance, if lower-caste youth overestimate caste-based socioeconomic disparities and caste-based discrimination, they might be more supportive of affirmative action policies that reduce these perceived disparities. Both lower and upper-caste youth might also have misperceptions about the proportion of India's population that is lower-caste. They might have normative beliefs about admission quotas for disadvantaged social groups being equivalent to that of the group's proportion in the population. If youth underestimate (overestimate) this number, they might prefer a lower (higher) threshold for the proportion of admission spots "reserved" or set aside for students from lower-caste backgrounds. Thus, misperceptions might influence support for redistributive policies such as affirmative action and moreover shape attitudes toward the beneficiaries of affirmative action.

In this study, I measure perceptions about caste among college-aged youth in India. Specifically, I measure their perceptions about India's population that is lower-caste, income disparities between upper-caste and different lower-caste groups (Scheduled Castes, Scheduled Tribe, and Other Backward Castes), and their perceptions about the education levels of India's lower-caste population. I further examine how these perceptions are related to preferences for affirmative action in higher education and attitudes toward caste. I find that while both upper and lower-caste youth have misperceptions, upper-caste youth misperceive to a greater extent. Upper-caste youth have larger misperceptions related to population and income disparities between caste groups. However, both upper and lower-caste groups equally misperceive education levels of India's lower-caste population.

I then examine whether correcting misperceptions through an experiment where I randomly assign respondents to receive facts related to population, income disparities, education level and research evidence about caste-based discrimination shifts their preferences for affirmative action and their attitudes toward caste. I find that while preferences for caste-based affirmative action remain unchanged, the intervention improves respondents' attitudes by 0.13SD toward beneficiaries of affirmative action and toward caste-based discrimination. Moreover, I find the effect on attitudes is larger (0.19SD) for upper-caste respondents. Finally, I re-run the analysis for a subset of respondents who passed a Factual Manipulation Check (FMC). FMCs are employed to identify respondents' attentiveness to the information provided by an experiment (Kane and Barabas, 2019). I find the treatment effect on attitudes is the largest (0.20SD) for respondents who are more attentive to the informational intervention.

This study is most closely related to a literature on misperceptions around inequality, affirmative action policies, and preferences for redistribution. There is a rich literature on misperceptions related to racial equity, social and economic mobility, and immigration in the United States and other developed countries (Alesina et al., 2018; Alesina and Stantcheva, 2020; Kraus et al., 2017). In the United States, for instance, individuals misperceive (overestimate) the extent of racial equity (Kraus et al., 2017). These studies further show that misperceptions shape attitudes toward different social groups and support for redistributive policies. Although the results are mixed, studies also show that attitudes toward disadvantaged social groups and redistribution can be changed with informational and narrative-based interventions (Alesina et al., 2023; Kuziemko et al., 2015). However, there is little research to my knowledge that captures misperceptions about inequality among college-aged students and proposes an intervention to counter these misperceptions.

The study is also related to the literature on affirmative action and attitudes toward caste among youth in higher education settings in India. Prior studies show that upper-caste students view lower-caste students—who are admitted through affirmative action—as less deserving and meritorious than upper-caste students (Deshpande, 2016; Deshpande and Ramachandran, 2019; Subramanian, 2015). Upper-caste students also perceive lower-caste students to be socially and economically privileged and less discriminated against (Desai and Dubey, 2012; Frisancho and Krishna, 2016). Lower-caste students, on the other hand, report that perceptions about their perceived privilege exacerbate the prevalence of negative attitudes toward them in higher education settings (Hasan and Bagde, 2013; Pathania and Tierney, 2018; Vijay and Nair, 2022). This study builds on our understanding of how (mis)perceptions shape negative attitudes toward caste and affirmative action and tests an informational intervention to mitigate them.

More broadly, I contribute to the literature on misperceptions between social groups and how these misperceptions might affect attitudes toward underrepresented groups and preferences for redistributive policies such as affirmative action. Affirmative action plays a pivotal role in redistributing resources toward underrepresented groups. Education systems in the United States, South Africa, Brazil, and India (among others) are grappling with how to improve inter-group relations in the context of redistributive policies that create beneficiaries and non-beneficiaries.

While I do not find any effect of my intervention on preferences for affirmative action, I find evidence that informational interventions can positively influence attitudes toward beneficiary groups in a setting where affirmative action policies are perceived as highly redistributive. Understanding the role that misperceptions play in shaping attitudes toward affirmative action and toward underrepresented beneficiary groups could help policymakers and administrators design policies that promote equity and diversity in higher education settings.

This paper is structured as follows: Section 2 provides background information on caste and affirmative action in higher education in India; Section 3 describes the data collection and the intervention; Section 4 provides descriptive results; Section 5 specifies the empirical strategy; Section 6 shows results of the randomized intervention; Section 7 discusses the implications of the findings and outlines next steps to take the study forward.

2.2 Background

The history of caste in India dates back to 1000 BCE (Thapar, 1999). Caste emerged as a form of social stratification based on a person's occupation. Occupational segregation and the inheritability of caste meant that there was very limited scope for intergenerational mobility. For certain groups, caste implied broader social and economic exclusion. At the heart of the exclusion was 'untouchability'—the practice of avoiding contact with those deemed of a lower caste. 'Untouchability' was tied to rules of ritual purity and pollution that governed upper-caste Hindu society. Thus, lower-caste groups were systemically excluded from many aspects of social and economic life for more than two millennia.

In 1947, India achieved independence from British colonialism and embarked on the creation of a modern state. The values of post-independence India were codified in its new constitution, which provided protections and preferences for previously disadvantaged caste groups. Specifically, the Indian constitution prohibited the practice of untouchability and put in provisions to reserve seats (quotas) for the most disadvantaged among lower-caste groups—the Scheduled Caste (SCs) and the Scheduled Tribes (STs)—in public universities, and government jobs. Affirmative action quotas in higher education and government jobs for SCs and STs were set at 15% and 7.5% respectively. In the early 1990s, affirmative action was extended to 'Other Backward Castes' (OBCs), and 27%

of seats were reserved for them in public universities and government jobs.[2] Collectively, Other Backward Classes (OBCs), Scheduled Castes (SCs), and Scheduled Tribes (STs) currently comprise 68% of India's current population and account for 49.5% of affirmative action quotas.[3]

India's public universities and colleges—which have admission quotas based on affirmative action—are also its most sought-after. Admissions to India's best public universities such as the Indian Institutes of Technology (IITs) is highly competitive with a cumulative acceptance rate of less than 2%.[4] In contrast to holistic college admissions in the United States, admission to India's public universities is often determined by a single high-stakes entrance exam. There are separate exams for different disciplines such as engineering, medicine, law, business management (MBA), etc. Based on the program students want to pursue, they take the entrance exam of their discipline. Universities then set entrance exam score "cut-offs" for admission to their institution and to specific degree programs.

For example, among the IITs, the more prestigious IITs such as IIT-Bombay have a higher admission score cut-off than newer (and less prestigious) IITs such as IIT-Indore. Even within a college, cut-offs differ by degree programs. For example, IIT-Bombay would have a higher rank or score cut-off for admission to its Computer Science degree program than to its Chemical Engineering degree program. The same process is applicable to state public colleges and universities as well. Though state universities are not as prestigious or selective as their national counterparts (such as the IITs), admission to them is still sought-after and more competitive than private institutions.

For admission into public universities and colleges, students are ranked only within their caste-categories quotas and thus compete with students who are of the same caste. Hence admission cut-offs are different for 'general' (which is the non-affirmative action category, comprised largely of upper-caste students), OBC, SC, and ST categories. Specifically, admission cut-offs are the highest in the 'general' or 'open' category, then in the OBC category, and are the lowest in the SC

---

[2]Caste is more granular than these categories suggest—according to some estimates India has more than 3,000 sub-castes (*jatis*) with varying levels of economic and social development.

[3]https://www.pewresearch.org/religion/2021/06/29/attitudes-about-caste/

[4]https://archive.nytimes.com/thechoice.blogs.nytimes.com/2011/10/14/indian-admissions/

and ST categories. On average, general-category students (who are mostly upper-caste) obtain higher test scores than OBC, SC, and ST students to gain admission to the same university and the same degree program. Admission to higher education institutions in India—especially public universities—is highly oversubscribed and the quota system potentially displaces upper-caste students who would have otherwise gained admission into a "better" public university or degree program in the absence of quotas.

Within this redistributive environment, opponents of caste-based affirmative action argue that socioeconomic disparities between upper and lower-caste groups have shrunk such that affirmative action in higher education is redundant. And if caste-based disparities are less salient, affirmative action would seem unfair and non-beneficiaries might feel resentful of the beneficiaries. In contrast, proponents of affirmative action argue that caste-based disparities are large and persistent, and Indian society is still rife with caste-based inequities and discrimination. Thus, there are two competing narratives about caste and affirmative action: one of large-scale convergence that has eliminated the need for affirmative action versus persistent caste-based disparities that validate affirmative action policies (Desai and Dubey, 2012).

From the perspective of convergence, Hnatkovska et al. (2013) use data from multiple rounds of the National Sample Survey (NSS) and show significant convergence since India's independence between upper and lower caste groups in education, income, and consumption—though they also point to gaps between these groups despite the convergence. The links between caste and occupation have also weakened post-independence (Béteille, 2012). Lower-caste groups have also made rapid strides in political representation (Jaffrelot, 2010). Part of the convergence can be attributed to India's affirmative action policies, which have been successful at targeting low-income lower-caste beneficiaries (Bagde et al., 2016; Bertrand et al., 2010; Prakash, 2020).

Despite this convergence, caste status is still an important determinant of educational and economic outcomes. Informal caste networks are powerful in India, and they determine access to education, high-status jobs, credit, and ownership of land and assets. Inequities might also manifest themselves in caste-based discrimination. Correspondence studies show caste-based discrimination in hiring (Siddique, 2011; Thorat and Attewell, 2007). In the context of higher education, a spate of student suicides on college campuses in India has been linked to caste-based discrimination against affirmative action beneficiaries.[5]

Thus, competing narratives—of convergence in caste-based disparities versus their persistence—shape beliefs about affirmative action among college-aged students. If these narratives are based on misperceptions, they might define youth's preference for affirmative action and attitudes toward its beneficiaries.

2.3 Data Collection and the Survey Experiment

2.3.1 Data Collection and Sample

I collected data for the study in August 2023. My sample is comprised of respondents who completed the entire survey. Only respondents over the age of 18 were allowed to take the survey. The survey was designed and administered via Qualtrics. Survey respondents were students enrolled at a private computer skills development institute located in Central India. The institute offers courses in computer programming and web and graphic design. Classes were conducted in-person across three centers in the city. Students at the institute were college students or recent college graduates.

Students were recruited via an email sent out by the institute. Students were told they could volunteer to take a short survey after class for which they would be paid Rs 50 ($0.60). They could take the survey on their own electronic devices. To minimize Hawthorne effects, students were not informed of the true purpose of the survey—that is, to capture misperceptions about caste. Instead, students were told that the survey was meant to learn about their experiences in college. Students who volunteered to take the survey stayed back after class. They were provided with the Qualtrics URL link for the survey. They were told that the survey was anonymous, and they should express their opinions honestly. They were also informed the survey was designed to take approximately 15-20 minutes to complete, and they should therefore take as much time as needed to respond. Respondents were paid in cash immediately after they completed the survey.

Table 2.1 describes the characteristics of respondents who completed the survey. It took respondents an average of 17.72 minutes to complete the survey. The sample consists of respondents

---

[5] https://www.outlookindia.com/national/caste-in-campus-from-rohith-vemula-to-darshan-solanki-cast

who are 32% female, 92% Hindu, 40% upper-caste, with an average age of 21.51 years. 45% of respondents have mothers who have at least completed schooling, 56% of the respondents are currently in college, and the average baseline preference of respondents for the percentage of seats that should be reserved in universities via affirmative action quotas is 44.98%. The respondents closely resemble the population of college students in India in terms of their religion and caste identification. For instance, the population of college students in India is 93% Hindu (92% Hindu in the sample) and 44% upper-caste (40% upper-caste in the sample).[6] However, the sample is less female (32%) compared to the population of college students in India who are women (49%).

2.3.2 Baseline Survey

The baseline survey was organized into two sections. The first section collected information about the background characteristics of respondents, such as their age, gender, caste, religion, education level, and socioeconomic background. The second section elicited their baseline perceptions about facts related to caste, their attitudes toward caste, and their preferences for affirmative action.

Respondents were asked about their perceptions related to caste on demographics, income, and education. Specifically, they were asked what was their best guess of India's population that is lower-caste (OBC + SC + ST combined), the average earning of a lower-caste person (OBC/SC/ST) for every Rs 100 that an upper-caste person earns, and their best guess for the percentage of lower-caste adults who have a college degree. Respondents were provided a user-friendly slider scale to respond. Next, respondents were asked to report their attitudes toward caste on a 5-point Likert scale ("Strongly Agree" to "Strongly Disagree"). Questions asked respondents about their perceptions of discrimination toward lower-caste students and whether lower-caste students who enter college through affirmative action should be provided financial aid. I used these attitude-related questions to construct an attitude index using factor analysis.

The respondents were then asked questions related to their preferences for affirmative action. The first set of questions related to preferences were also on a Likert scale and asked respondents their views on the fairness of caste-based affirmative action, whether caste-based affirmative action

[6] Estimates of the population of college students in India are from the All India Survey of Higher Education (AISHE) 2020-21: https://aishe.gov.in/aishe/viewDocument.action?documentId=352

in higher education reduced inequality, and whether there should be fewer seats reserved via caste-based affirmative action. I constructed an index for preferences using these preferences-related survey questions. Finally, respondents were asked to indicate their baseline preference for the percentage of seats (quota) that should be reserved for caste-based affirmative action in higher education. This question was presented on a slider scale of 0-100%, with 0 representing no seats should be reserved for candidates from lower-caste groups and 100 indicating all seats should be reserved for them.

2.3.3 Intervention and Endline Survey

After completing the baseline survey, respondents were randomly assigned (automatically via Qualtrics software) to either a placebo or to the treatment. A balance table in the appendix (Table 2.9) shows that respondents in the treatment and control groups are balanced on observables, and the randomization was successful. Those assigned to the treatment arm were first provided facts about India's caste composition, income disparities between upper and lower-caste groups, and the education level of lower-caste adults.[7] The information provided facts directly related to the caste-related perception questions respondents were asked in the baseline. The facts were provided in the form of colorful figures and graphics.

Respondents were then directed to watch a short video ($\sim$90 seconds) on research evidence regarding caste-based discrimination in the job market. Specifically, the video showed research evidence using correspondence/audit studies conducted in India that show that the upper and lower-caste applicants with identical CVs were treated differently. Correspondence studies were chosen as part of the research evidence treatment based on pilot work that indicated that these studies were intuitive for students to understand. The videos and graphics were professionally produced in the format of an Instagram reel to help keep students engaged.

Those assigned to the placebo group were provided information unrelated to the intervention. The placebo was designed to take approximately the same time as the treatment. The placebo also followed a structure similar to the treatment—some information followed by a short (though

[7] Estimates of India's population that is lower-caste were obtained from Pew Research Center's survey about caste in India (Pew, 2021). Facts related to income and education levels were constructed using National Sample Survey (NSS) data.

unrelated to the treatment) video.

Immediately after the interventions, both treatment and placebo groups were administered a short endline survey. The survey again asked respondents about their attitudes toward caste and their preferences for affirmative action (both on a Likert scale and in percentage of reservation quota). Finally, respondents were asked a question related to the information provided in the interventions. Based on the arms to which the respondents were assigned (placebo or treatment), they were asked a different question. These questions—as part of a Factual Manipulation Check (FMC)—were meant to test respondents' attentiveness to the intervention.

2.4 Perceptions about Caste

Table 2.2 Panel A shows the difference in perceived (reported by respondents) and actual facts related to a) India's population that is lower-caste (OBC + SC + ST combined) b) income disparities between upper-caste and various lower-caste groups and c) the percentage of lower-caste adults in India who have a college degree. To estimate respondents' perceptions of income disparity, the prompt asked, "For every Rs 100 that an average upper-caste person earns, what is your best guess for the average earning of an average OBC/SC/ST person? (Rs 100 would mean earning equal to an upper-caste person)". Respondents were provided a separate slider scale for each lower-caste group to indicate their best guess for that group.

Table 2.2, Panel A shows results for the full sample. The average survey respondent underestimates India's population that is lower-caste by approximately 4 percentage points (pp). Misperceptions in the income and education categories are more striking. The average respondent underestimates the income disparity between upper-castes and OBCs by 14.45 pp, between upper-caste and SCs by 17.10 pp, and between upper-caste and ST by 22.23 pp. In general, the average respondent also underestimates income disparities *between* lower-caste groups. Finally, the average respondent overestimates the level of affirmative action in higher education in terms of the percentage of seats that are reserved or in public colleges for students from lower-caste groups by 3.3 pp.

Table 2.2, Panel B and Panel C break misperceptions down by respondents' caste. The results show that, on average, there are larger differences in perceived and actual facts among upper-caste respondents compared to their lower-caste counterparts. Upper-caste respondents underestimate India's population that is lower-caste by 6.6 pp and underestimate income disparities between upper-castes and OBCs, SCs, and STs by 18.18 pp, 21.63 pp, and 26.11 pp, respectively. However, both upper and lower-caste respondents perceive 60% of lower-caste adults in India have a college degree when 8% actually do—a discrepancy of 52 pp. Moreover, while both upper and lower-caste respondents overestimate the percentage of seats reserved for lower-caste groups in public higher education institutions, upper-caste respondents overestimate it marginally by 2pp (although this misperception is only significant at the 10% level) and lower-caste respondents overestimate it to a greater extent by 4.86 pp. Figure 2.1 presents a comparison of misperceptions for upper-caste and lower-caste respondents with 95% confidence interval around the average misperceptions. The differences in misperception are larger for upper-caste respondents on questions related to population and income disparities between upper-castes and OBCs and SCs, but statistically not different for income disparities between upper-castes and STs and perceptions of the education levels of lower-caste adults.

The descriptive part of the study has four key findings. First, both upper-caste and lower-caste respondents *underestimate* India's population that is lower-caste, *underestimate* income disparities between upper and lower-caste groups, and *overestimate* the education levels of lower-caste adults. However, upper-caste respondents tend to have larger misperceptions regarding India's population and caste-based income disparities. Second, both groups underestimate income disparities *between* lower-caste groups—thus they underestimate disparities between OBCs and SC, OBCs and STs, and SCs and STs. Third, both upper-caste and lower-caste respondents have striking misperceptions regarding the percentage of lower-caste adults in India who have a college degree. Fourth, upper-caste respondents have more accurate perceptions of the percentage of seats reserved in higher education via caste-based affirmative action compared to their lower-caste counterparts.

2.5 Empirical Strategy

To estimate the results of my survey experiment, I use the following empirical specification:

$$Y_i = \alpha + \ \gamma_1 D_{1i} + \beta X_i + \epsilon_i$$

Y_i is the outcome of interest measured for student i; D_{1i} is a binary indicator for treatment status, and X_i is the vector of baseline characteristics, and ϵ_i is the error term. Baseline characteristics include female (1/0), upper-caste (1/0), Hindu (1/0), age, whether the respondent is currently in college (1/0), whether the respondent's mother has completed schooling (1/0), whether the respondent has a bachelor's degree (1/0), and outcomes measured at the baseline.

2.6 Results

2.6.1 Main Results

In Table 2.3, I present the main results of the experiment.[8] I have three primary outcomes: preference for the percentage of seats that should be reserved in universities via caste-based affirmative action quotas; a preference score (Z-scored) based on four questions in the survey that capture respondents' support for affirmative action policies; and an attitudes score (Z-scored) based on four questions in the survey that capture respondents' attitudes toward lower-caste affirmative action beneficiaries and toward caste-based discrimination.

The estimated effect of the treatment on the preference for the percentage of seats reserved via affirmative action quotas is small and negative (-0.77pp) and that on the preference score is small and positive (0.008SD) but the standard errors are large, and the estimates are imprecise. Thus, I do not find an effect of being assigned to the treatment on either of the preference outcomes. However, I find that my treatment has an effect on attitudes scores—respondents who are assigned to the treatment report positive attitudes toward lower-caste beneficiaries of affirmative action and toward caste-based discrimination. Specifically, treated respondents scored 0.13SD (p = 0.014**) higher on the attitudes compared to those in the control group. I also adjust for multiple hypothesis testing (Table 2.3) and find that my results remain unchanged.

[8] After adding the full-set of controls, I lose 52 observations due to missingness. However, my results are substantively the same with and without controls—thus, for brevity, I only present results with controls in the main paper.

2.6.2 Factual Manipulation Check (FMC) Sample Results

In the endline, I ask respondents in both the treatment and placebo group a question related to their informational interventions to test for attentiveness. 64% of the respondents correctly answered the question and passed the Factual Manipulation Check (FMC). Table 2.4 presents results for this subsample of respondents. While I still do not find an effect of the treatment on outcomes related to preferences for affirmative action, I find that the treatment has a strong and positive effect on attitudes. Specifically, treated respondents score 0.20SD ($p < 0.01$***) higher on attitudes compared to those in the control group. The precision of my estimates also improves. I adjust my results for multiple hypothesis testing and the significance level remains the same. Results for the subsample of respondents who passed the FMC suggest the treatment had a stronger effect on those who were more attentive to the information provided by the intervention.

2.6.3 Upper-caste and Lower-caste Sample Results

In Section 4, I show that, on average, upper-caste respondents report larger misperceptions related to caste. If correcting misperceptions is an important mechanism that improves respondents' attitudes toward caste, then the treatment effect should be larger, on average, for treated upper-caste respondents. In Table 2.5, I present results for the subsample upper-caste respondents. As with the full sample, while I do not find an effect of the treatment on outcomes related to preferences for affirmative action, I find a positive effect on attitude scores. Specifically, treated upper-caste respondents scored 0.19SD ($p = 0.018$**) higher on attitudes compared to upper-caste respondents in the control group.[9] In Table 2.6, I also present results for the subsample lower-caste respondents. I find a weaker and less significant effect of the intervention on the attitudes of lower-caste respondents. Specifically, treated lower-caste respondents scored 0.13SD ($p = 0.019$*) higher on attitudes compared to lower-caste respondents in the control group. However, this effect is no longer significant after adjusting for multiple hypothesis testing.

In Table 2.7, I present results for the subsample of upper-caste respondents who passed the FMC. In this subsample, treated upper-caste respondents scored 0.26SD ($p < 0.01$***) higher on

[9] After adjusting for multiple hypothesis testing using three separate tests, the significance level on the attitudes score falls to 10% on two of the three tests.

the attitudes compared to upper-caste respondents in the control group. The results remain unchanged after adjusting for multiple hypothesis testing.

2.6.4 Robustness Check

Respondents who take too little or too much time to complete the survey might be systemically different from the average respondent. For instance, spending very little time on the survey might indicate rushed responses, and spending too much time might indicate issues with comprehension and attentiveness. Following Alesina et al. (2023), I trim the top and bottom 2% of the respondents in the distribution of the duration to complete the survey and re-run my analysis. Trimming does not change my main results. The treatment effect on the set of preference outcomes remains small and insignificant, whereas the effect on the attitudes score—0.124SD (p=0.026**)—is similar to that for the full sample (Table 2.11).

2.7 Mechanisms

Since the effect of the intervention is limited to the change is attitude scores, I further investigate which component of attitudes is driving the change. The attitudes index (score) is comprised of four separate components: whether lower-caste students are discriminated in college, whether lower-caste students are discriminated in the job market, whether lower-caste students who are admitted through affirmative action should be offered scholarship and whether a majority of lower-caste college students come from rich families. I standardized each component and regress each separately on the treatment indictor with control variables for the full sample. The results of the analysis are presented in Table 2.8. I find the two discrimination-related component are driving the change in attitudes. Specifically, treated respondents score 0.16 SD (p=0.006***) on whether lower-caste students are discriminated in college and 0.15 SD (p=0.015**) on whether lower-caste students are discriminated in the job market. Treated respondents also score higher on whether lower-caste students who are admitted through affirmative action should be offered scholarships by 0.07 SD (p=0.30), though the effect is not statistically significant.

I also constructed a misperception index based on respondents' misperceptions about population level, income disparities, and education level of lower-caste groups. I find a negative relationship between baseline attitudes and perception scores and the size of misperceptions. Specifically, I find that a 1SD higher score on the misperception index (larger misperceptions) is associated with -0.14SD (p=0.000***) worse attitudes in the baseline and a lower preference score on the preference index by -0.18SD (p=0.000***). This suggests that those with larger misperceptions about caste tend to have less favorable attitudes toward lower-caste beneficiaries of affirmative action and lower preferences for affirmative action.

2.8 Discussion and Conclusion

In this paper, I document misperceptions about caste among upper-caste and lower-caste youth in India. I provide evidence that both upper-caste and lower-caste youth have striking misperceptions about caste-based income disparities and the education levels of lower-caste adults. And, on average, upper-caste youth have larger misperceptions. These misperceptions occur in the broader context of affirmative action in higher education in India, where lower-caste youth are beneficiaries and upper-caste youth are non-beneficiaries. My result suggests that misperceptions related to caste might be shaping and misinforming conversations about caste and affirmative action in higher education settings in India.

I also document evidence that correcting misperceptions can alter attitudes toward lower-caste beneficiaries of affirmative action and toward caste-based discrimination, but does not alter preferences for affirmative action. The evidence aligns with previous studies that use informational evidence to correct misperceptions toward immigrants, political partisans, and racial groups, and find changes in attitudes but a small or no change in preferences for redistribution (Grigorieff et al., 2020; Kuziemko et al., 2015; Mernyk et al., 2022). My findings add to the evidence that preferences might be less malleable than attitudes.

It is important to note that I conducted my study in an environment that is heterogeneous with respect to caste—40% of my sample is upper-caste, and 60% is lower-caste. Such heterogeneous environments might encourage greater interaction and contact across caste lines and mitigate caste-related misperceptions. Misperceptions might be larger in homogenous settings (such as elite private universities in India) where upper-caste students in particular have less contact with lower-caste students. The findings of this study, which suggest that upper-caste youth—who have larger misperceptions in baseline—are more responsive to the informational intervention, could be even more relevant in homogenous higher education settings in India.

This study has three potential limitations. First, I am only able to capture the effects of the intervention on attitudes in the short-run. It is unclear whether the effect of correcting misperceptions on attitudes persists in the long-run. Second, this study only captures self-reported attitudes and preferences, which might be vulnerable to biases and measurement error. Third, the study's intervention provides information in the form of facts *and* research evidence. I cannot entirely disentangle which aspect of the intervention—fact or research evidence—is more effective in improving attitudes toward caste.

Overall, I contribute to our understanding of misperceptions among beneficiaries and non-beneficiaries in environments that have redistributive policies such as affirmative action. This study further provides suggestive evidence that attitudes toward affirmative action beneficiaries can be improved through a low-cost, light-touch informational intervention. This evidence can be used in higher education settings in India and elsewhere to better inform conversations about affirmative action and to promote equity and diversity.

Tables

Table 2.1: *Sample Characteristics*

	Mean	SD
Duration complete survey	17.72	6.73
Female	0.32	0.47
Hindu	0.92	0.26
Age	21.51	2.93
Currently in college	0.56	0.50
Uppercaste	0.40	0.49
Baseline preference affirmative action	44.98	26.18
Mother completed schooling	0.45	0.50
Observations	774	

Notes: This table presents descriptive statistics for the sample of 774 respondents. Duration complete survey is measured in minutes. Female, Hindu, Currently in college, Uppercaste, and Mother completed schooling are binary variables (1=yes, 0=no). Baseline preference affirmative action is measured on a 0-100 scale, where higher values indicate stronger support for affirmative action policies. The sample consists of college-aged youth from institutions in India.

Table 2.2: *Misperceptions Related to Population, Income, and Education*

	Perceived	Actual	Diff	p-value
Panel A: Full Sample				
Perceived vs actual % of India's population lower-caste	64.02	68.00	-3.98	0.00
Perceived vs actual income difference: upper-caste vs OBC	78.45	64.00	14.45	0.00
Perceived vs actual income difference: upper-caste vs SC	74.10	57.00	17.10	0.00
Perceived vs actual income difference: upper-caste vs ST	72.33	50.00	22.33	0.00
Perceived vs actual % lowercaste adults w/ college degree	59.82	8.00	51.82	0.00
Perceived vs actual % of seats reserved for lowercaste students	53.30	50.00	3.30	0.00
Panel B: Upper-Caste Sample				
Perceived vs actual % of India's population lower-caste	61.39	68.00	-6.61	0.00
Perceived vs actual income difference: upper-caste vs OBC	82.18	64.00	18.18	0.00
Perceived vs actual income difference: upper-caste vs SC	78.63	57.00	21.63	0.00
Perceived vs actual income difference: upper-caste vs ST	76.11	50.00	26.11	0.00
Perceived vs actual % lowercaste adults w/ college degree	59.94	8.00	51.94	0.00
Perceived vs actual % of seats reserved for lowercaste students	52.01	50.00	2.01	0.09
Panel C: Lower-Caste Sample				
Perceived vs actual % of India's population lower-caste	65.99	68.00	-2.01	0.00
Perceived vs actual income difference: upper-caste vs OBC	75.80	64.00	11.80	0.00
Perceived vs actual income difference: upper-caste vs SC	70.72	57.00	13.72	0.00
Perceived vs actual income difference: upper-caste vs ST	70.08	50.00	20.08	0.00
Perceived vs actual % lowercaste adults w/ college degree	59.97	8.00	51.97	0.00
Perceived vs actual % of seats reserved for lowercaste students	54.86	50.00	4.86	0.00

Notes: To capture perceptions about disparities in income, the survey asked respondents their best guess about how much an average lower-caste person (OBC, SC, and ST) earned for every Rs 100 that an average upper-caste person earned.

Table 2.3: *Treatment Effect on Preferences and Attitudes, Full Sample*

	(1) Preference % Quota	(2) Preferences Index	(3) Attitudes Index
Treatment	-0.775	0.008	0.135^{**}
	(1.266)	(0.063)	(0.055)
N	722	722	722
R^2	0.533	0.272	0.435
Bonferroni Holm pvalues	1.000	1.000	0.042
Westfall Young pvalues	0.781	0.888	0.038
Sidak Holm pvalues	0.789	0.901	0.042

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: Robust standard errors in parentheses. Treatment is a dummy variable indicating random assignment to the informational arm that provides facts and research evidence about caste. Preference % Quota refers to respondents' preference for % of seats that should be reserved in public universities for caste-based quotas. Attitudes Index and Preferences Index outcomes are in z-scores with mean = 0 and SD = 1. All specifications include a full set of controls.

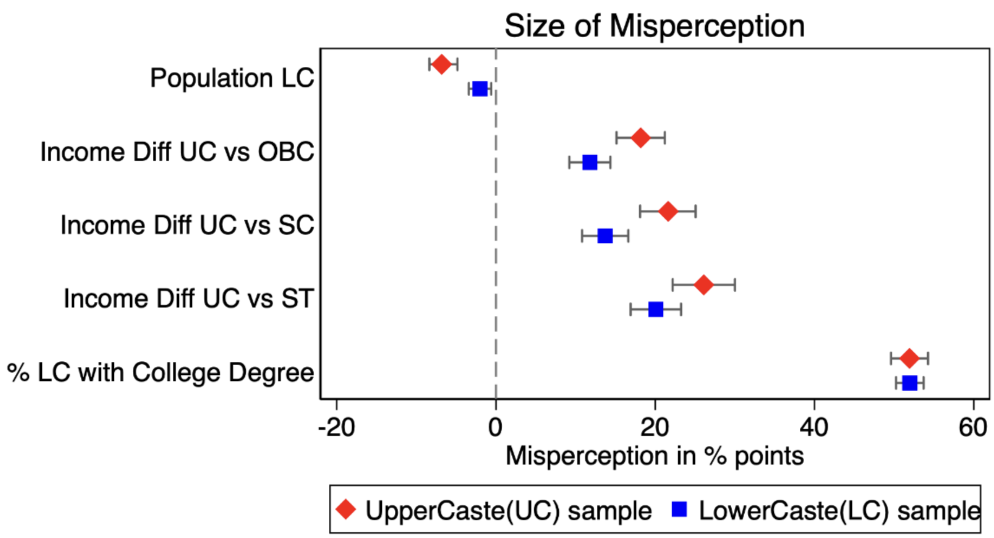


Figure 2.1: *Misperceptions by Caste (Difference in Perceived minus Actual)*

Notes: This figure shows the difference between reported/perceived vs actual estimates of India's population that is lower-caste (OBC + SC + ST), income disparities between upper-caste and each lower-caste subgroup (OBC/SC/ST), and the percentage of lower-caste adults who have a college degree. The figures show these estimates separately for upper-caste and lower-caste respondents. Outcomes are on the y-axis and the size of misperception in percentage points is on the x-axis.

Table 2.4: *Treatment Effect, FMC Sample*

	(1) Preference % Quota	(2) Preferences Index	(3) Attitudes Index
Treatment	−1.720	0.048	0.198***
	(1.684)	(0.074)	(0.064)
N	470	470	470
R^2	0.503	0.309	0.474
Bonferroni Holm pvalues	0.615	0.615	0.006
Westfall Young pvalues	0.516	0.531	0.008
Sidak Holm pvalues	0.521	0.521	0.006

Robust Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.5: *Treatment Effect, Upper-Caste Sample*

	(1) Preference % Quota	(2) Preferences Index	(3) Attitudes Index
Treatment	-0.023	0.074	0.191^{**}
	(2.092)	(0.099)	(0.080)
N	315	315	315
R^2	0.511	0.266	0.469
Bonferroni Holm pvalues	0.991	0.901	0.054
Westfall Young pvalues	0.994	0.692	0.046
Sidak Holm pvalues	0.991	0.698	0.053

Robust standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.6: *Treatment Effect, Lower-caste Sample*

	(1) Preference % Quota	(2) Preferences Index	(3) Attitudes Index
Treatment	-0.971	-0.009	0.132^{*}
	(1.509)	(0.078)	(0.078)
N	453	453	453
R^2	0.508	0.310	0.317
Bonferroni Holm pvalues	1.000	1.000	0.271
Westfall Young pvalues	0.765	0.911	0.219
Sidak Holm pvalues	0.770	0.905	0.247

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.7: *Treatment Effect, Upper-Caste FMC Sample*

	(1) Preference % Quota	(2) Preferences Index	(3) Attitudes Index
Treatment	-0.557	0.053	0.260^{***}
	(2.524)	(0.119)	(0.084)
N	221	221	221
R^2	0.531	0.293	0.563
Bonferroni Holm pvalues	1.000	1.000	0.007
Westfall Young pvalues	0.879	0.879	0.004
Sidak Holm pvalues	0.882	0.882	0.007

Robust standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.8: *Treatment Effect by Components of Attitudes Index, Full Sample*

	(1) Discrimination College	(2) Lowercaste Scholarships	(3) Discrimination Jobs	(4) Lowercaste Rich
Treatment	0.155***	0.070	0.146**	0.030
	(0.006)	(0.330)	(0.015)	(0.680)
Observations	722	722	722	722
R^2	0.422	0.060	0.338	0.034

p-values in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Chapter 2 Appendix

Table 2.9: *Balance on Pretreatment Covariates*

Variable	Control Mean (SD)	Treatment Mean (SD)	Difference	(p-value)
Female	0.31	0.33	0.02	
	(0.46)	(0.47)		(0.57)
Hindu	0.92	0.93	0.01	
	(0.27)	(0.26)		(0.73)
Bachelors degree	0.78	0.80	0.02	
	(0.42)	(0.40)		(0.51)
Age	21.69	21.24	-0.35^{*}	
	(3.16)	(2.67)		(0.09)
Currently in college	0.55	0.57	0.02	
	(0.50)	(0.50)		(0.68)
Baseline attitude index	−0.04	0.04	0.08	
	(0.98)	(1.01)		(0.28)
Baseline preference index	−0.01	0.01	0.02	
	(0.99)	(1.02)		(0.81)
Uppercaste	0.41	0.40	−0.00	
	(0.49)	(0.49)		(0.94)
Baseline preference affirmative action	45.14	41.13	−1.62	
	(26.79)	(25.60)		(0.95)
Mother completed schooling	0.45	0.44	−0.01	
	(0.50)	(0.50)		(0.76)
Observations	379	385	774	

Notes: Columns (1) and (2) display the estimated means of pretreatment covariates for the control and treatment groups, respectively. Standard deviations are in parentheses. Column (3) displays the difference in means between the two groups, with p-values in parentheses. Joint F-test p-value for all covariates: 0.9217.

Table 2.10: *Balance on Pretreatment Covariates, Dropping Missing Values*

Variable	Control Mean (SD)	Treatment Mean (SD)	Difference	(p-value)
Female	0.31	0.35	0.02	
	(0.46)	(0.47)		(0.53)
Hindu	0.93	0.92	−0.00	
	(0.26)	(0.27)		(0.91)
Bachelors degree	0.78	0.80	0.01	
	(0.41)	(0.40)		(0.62)
Age	21.68	21.33	−0.35	
	(3.19)	(2.66)		(0.11)
Currently in college	0.54	0.57	0.03	
	(0.50)	(0.50)		(0.41)
Baseline attitude index	0.05	0.03	0.08	
	(0.95)	(1.04)		(0.28)
Baseline preference index	−0.02	−0.02	0.01	
	(0.97)	(1.02)		(0.92)
Uppercaste	0.41	0.40	−0.00	
	(0.49)	(0.49)		(0.97)
Baseline preference affirmative action	44.87	44.62	−0.25	
	(26.73)	(25.62)		(0.90)
Mother completed schooling	0.45	0.44	−0.01	
	(0.50)	(0.50)		(0.70)
Observations	353	369	722	

Notes: Columns (1) and (2) display the estimated means of pretreatment covariates for the control and treatment groups, respectively. Standard deviations are in parentheses. Column (3) displays the difference in means between the two groups, with p-values in parentheses.

Table 2.11: *Treatment Effect on Preferences and Attitudes, Robust*

	(1) Preference % Quota	(2) Attitudes Index	(3) Preferences Index
Treatment	−0.633	0.124**	0.002
	(1.287)	(0.056)	(0.063)
Observations	703	703	703
R^2	0.535	0.433	0.275

Robust standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Chapter 3

Women in STEM: Access, Selection, and Confidence in a Training Program for Women Engineers in India

Joint work with Saloni Gupta at Brown University

3.1 Introduction

The under-representation of women in STEM (science, technology, engineering, and mathematics) constitutes a significant barrier to fostering a robust and inclusive STEM workforce. This gap stems from cultural, institutional, and psychological factors, with negative norms and stereotypes hindering women's success (Ceci and Williams, 2011; Kahn and Ginther, 2018). In developing countries, these challenges are aggravated by gendered resource allocation and mobility constraints, which limit women's access to enrichment opportunities and high-quality technical education. Can we identify programs that address some or all of these challenges for women in developing countries?

In this paper, we conduct a descriptive analysis of a program designed to address these barriers to women's STEM education in developing countries. The program is a 24-month immersive, live online experiential learning program administered by an India-based ed-tech company in partnership with a major global technology firm. It selects, trains, and nurtures first-year women engineering students to become software engineers prepared for top technology jobs, combining technical coursework, team-based projects, soft-skills training, and exposure to alumni and industry mentors. The program has been active since 2019 and, at the time of this analysis, has admitted six cohorts of women engineering students drawn from colleges across India, with the applicant pool growing steadily over time. Our objective is to assess the accessibility of the program for women of low-income backgrounds and rural areas, and thus to understand its targeting and geographic coverage.

Across the first six cohorts since 2019, we find that the program attracts a broad applicant pool: 55–60% of applicants come from rural areas, and 60–85% come from low-income families. While the selection process skews toward more urban and higher-income candidates, the program nonetheless maintains meaningful access for disadvantaged populations: 28–56% of selected participants come from low-income families, 53–77% attend non-elite colleges, and participants are drawn from more than 20 states (71% of states) across India. Given the program's highly competitive acceptance rate (0.8–1.4%), these patterns suggest the program provides access to high-quality technical training for women who might otherwise lack such opportunities.

The program itself overcomes barriers faced by women in STEM in several ways. First, the program recruits only women to address the difficulty women face in entering and thriving in STEM fields traditionally dominated by men. Research suggests that the presence of same-gender peers positively impacts cognitive outcomes and self-confidence, benefiting women in both schooling and higher education (Eisenkopf et al., 2015; Lavy and Schlosser, 2011; Oosterbeek and van Ewijk, 2014). Specifically in the context of higher education, having a greater number of women peers correlates with increased retention in STEM majors, on-time graduation, and improved labor market outcomes (Bostwick and Weinberg, 2022; Booth et al., 2018; Karpowitz et al., 2024; Mouganie and Wang, 2020). This would seem to suggest that this program, which creates women-only spaces, could positively impact participants' outcomes.

Mobility barriers, especially in rural settings, restrict freedom of movement for women and undermines their participation in social and educational activities (Kantor, 2002). Since this program is fully online, it removes these mobility constraints and allows women to access this training program from home or college and participate in online mentoring and networking events. This increased access can, in turn, positively impact their career advancement and opportunities in the STEM field, potentially bridging gender gaps in educational opportunity. This characteristic holds promise in democratizing access to education by eliminating barriers associated with physical infrastructure, geographic location, or resource constraints, thus facilitating greater participation in learning opportunities (Acemoglu et al., 2014).

Finally, the program incorporates self-directed learning, an instructional approach that asks participants to take initiative and responsibility for their own learning. By engaging in team-based problem-solving of engineering challenges, participants set their own goals, identify learning resources, and evaluate their progress. Pedagogies of this kind have been linked to improvements in problem-solving skills and in self-efficacy in technical domains (Eccles and Wigfield, 2002).

As noted earlier, the primary objective of this paper is to analyze this national online learning initiative and to better understand whether it is successful in targeting disadvantaged women engineering students. Whether selective education and training programs reach the disadvantaged students they aim to serve is a long-standing concern. Even when generous financial support is available, many low-income high-achievers do not apply, often because of information frictions, complexity, or geographic distance from program outreach (Hoxby and Avery, 2013; Dynarski et al., 2021). To accomplish this, we examine the demographic characteristics of students who apply to and get selected into the program. We used organizational data that span the past five cohorts, along with a recently conducted baseline survey in February 2024 among participants recruited in the sixth and latest cohort.

An additional objective of this paper is to introduce a novel interview-based measure to assess the confidence of women engineers in developing country settings. Interviews play an important role in selection processes in the educational and labor markets. They often determine access to scholarships, selective programs, and employment opportunities. One factor that can shape success in such settings is how external evaluators perceive the confidence of a candidate. Women may be particularly vulnerable to penalties associated with being perceived as less confident, especially in male-dominated STEM fields (Alan et al., 2018).

In this study, we leverage a unique feature of the program's baseline data collection for the sixth cohort (2024-2026). As part of the selection process, the candidates participated in structured interviews in which the interviewers rated the participant's confidence on multiple dimensions. These interviewer ratings capture external perceptions of confidence rather than candidate self-beliefs. We also capture candidates self-reported confidence through the General Self-Efficacy scale administered during the baseline. We find that while low-income and higher-income students show identical levels of self-reported confidence, interviewers perceive low-income students as significantly less confident. This "perception gap" is nearly twice as large for students from economically disadvantaged backgrounds, suggesting that they may face a penalty in competitive selection settings despite having the same internal confidence as their higher-income peers. Our ongoing randomized evaluation will assess whether participation in the program can reduce this perception gap.

The remainder of the paper proceeds as follows. Section 2 describes the training program and its key components. Section 3 outlines the data sources used in the analysis, including administrative records, baseline survey data, and interviewer-rated measures of confidence collected during the program's selection process. Section 4 presents descriptive results on the composition of applicants and participants, access and selectivity patterns, baseline characteristics of women admitted to the program, and our analysis of the confidence perception gap by socioeconomic background. Section 5 concludes with a discussion of how these findings inform our understanding of the barriers women face in STEM and outlines our ongoing randomized evaluation.

3.2 Intervention

Despite comprising 43% of STEM graduates in India, women remain significantly underrepresented in science and technology careers. For instance, only 27% are represented in the STEM workforce and only 14% in C-suite roles in STEM.[1] Research suggests that women leave STEM fields due to a combination of workplace climate, lack of mentorship, and work-life balance challenges (Hunt, 2016). This emphasizes the importance of nurturing talent in STEM so that more women can transition from higher education to the job market and also to senior roles within industry.

This paper examines a women's engineering program that operates nationwide in India and aims to nurture talent. In a future paper, we will evaluate the impact of the program on improving women's outcomes. The program is facilitated by an undisclosed organization located in India, and we are obligated to maintain confidentiality regarding its identity. This program is a 24-month online learning experiential program that aims to select, train, and nurture first-year engineers for top engineering jobs. Selected participants are awarded a full scholarship that covers the entire program fee, along with a one-time scholarship to help pay for personal expenses.

[1]See EY, *Breaking the Code: The Rise of Women in India's STEM Landscape* (2026) and Anuradha Mascarenhas, "Women, STEM careers, and a more receptive industry," *The Hindu*, July 2024, `https://www.thehindu.com/opinion/op-ed/women-stem-careers-and-a-more-receptive-industry/article69811907.ece`.

The organization typically opens applications in January for first-year women engineers from colleges across India, with selections announced in March. The program, which spans 24 months with three-hour weekend sessions, begins in the summer after the first year of college and ends in the first semester of the third year. The program also exposes participants to alumni with successful STEM careers, serving as female role models. Research demonstrates that exposure to female role models can reduce gender stereotypes and increase women's participation in STEM (Breda et al., 2023; Porter and Serra, 2020; Ahmed et al., 2024).

The program has been active since 2019 and has successfully launched six cohorts of students. During the last six cohorts, the number of applications has increased steadily, making the program more competitive, as shown in Table 3.3. We have completed the analysis of the first three cohorts who have transitioned from their engineering degrees to the workforce, providing us with valuable insights into their labor market outcomes.

We also incorporate findings from a baseline survey administered to the most recent cohort admitted in February 2024. This cohort is part of an RCT evaluation of the program that will conclude in April 2026. In addition to the baseline survey, we collected and evaluated a novel metric to assess students' confidence. We examine the confidence levels of students as evaluated by interviewers during the recruitment process. This measure will serve as an important indicator for monitoring students' confidence over time.

In terms of the specifics of the program, it is designed and implemented by an ed-tech organization with substantial experience in training engineering graduates in India. The program itself has both technical and soft skills components. The curriculum is distributed as shown in Table 3.1.

The program consists of four courses, each 120 hours long. In total, the program involves 720 hours of coursework. This is in addition to the coursework that students take independently in their colleges. Thus, this program requires a significant time commitment from the participating students. Three courses focus on technical and coding skills, with both individual- and team-based components. Similar programs focused on coding skills for women have shown positive effects on skills and career aspirations (Carlana and Fort, 2022; Basiglio et al., 2024). Although technical skills are the core component of the first three courses, they have soft skills components embedded in the coursework. For example, students frequently work in teams and on projects that require

Table 3.1: *Program Overview and Course Objectives*

Component	**Objectives**	**Hours**
Course 1: Building CLI applications, using Python. Git, team work	Write clean, functionally decomposed Python code. Understand and use imperative, functional and object-oriented paradigms. Build classes and simple class hierarchies. Use git effectively. Work in teams	120
Project 1: Team Project Work	Apply knowledge gained in Course 1 to real-world projects	60
Course 2: Building simple web applications using JS/HTML/CSS and db	Build web applications	120
Project 2: Team Project Work	Apply knowledge gained in Course 2 to real-world projects	60
Course 3: Advanced Data Structures and algorithms	Develop an understanding of basic data structures and algorithms. Make students taste the joy of problem-solving. Develop an attitude for writing optimized code	120
Project 3: Team Capstone Project	Apply knowledge gained in Courses 1-3 to real-world projects	12
Course 4: Soft Skills	Enhance personal accountability and professionalism. Improve professional communication skills. Foster critical thinking abilities. Prepare for job interviews	120
Total	4 Courses, 3 Projects	**720**

them to collaborate effectively. The program also has a course (Course 4) dedicated exclusively to developing soft skills among the program's participants. Within the soft skills course, the program offers five modules. Each module is carefully designed to equip students with skills that the training organization (e.g., in its survey of industry recruiters) has identified as crucial to students' success in the labor market. These modules are outlined in Table 3.2.

Module 1, "Personal Accountability," covers personal goal setting, time management, and taking ownership of one's professional development. Module 2, "Professional Communication," develops skills for conveying ideas in professional settings, including delivering presentations (Present Right), developing emotional intelligence (IQ to EQ), and engaging in workplace in-

Table 3.2: *Modules Offered in the Soft Skills Course*

Module 1: Personal Accountability	Module 2: Professional Communication
Personal Accountability KASH Model Goal Setting Time Management Write Right	Present Right IQ to EQ Interact, Interpret, Respond Crossing the Communication Chasm Concise, Cogent Communication
Module 3: Technical Expositions and Discussions	**Module 4: Critical Thinking**
Organization Key Points Differing Opinions Logical Conclusions	Introduction to Critical Thinking Non-Linear Thinking Logical Thinking Characteristics of a Critical Thinker Problem-Solving
Module 5: Interview Readiness	
Resume Building Cover Letter Writing Interview Readiness in the Age of AI Telling Your Story LinkedIn Profile Building - Building Professional Brand	

teractions (Interact, Interpret, Respond). Module 3, "Technical Expositions and Discussions," focuses on technical communication: organizing information (Organization), distilling key points (Key Points), navigating differing opinions (Differing Opinions), and drawing logical conclusions (Logical Conclusions). Module 4, "Critical Thinking," introduces different thinking patterns (Introduction to Critical Thinking), non-linear problem-solving (Non-Linear Thinking), logical reasoning (Logical Thinking), and practical problem-solving exercises (Problem-Solving). Module 5, "Interview Readiness," prepares participants for the job market through resume building, cover letter writing, interview preparation in the age of AI, articulating one's professional story, and building a LinkedIn presence.

Because of the structured modules outlined above, one of the research questions we will be able to ask is how these modules contribute to the development of specific soft skills among participants. In the endline scheduled for March 2026, we map specific outcomes to each module and measure the overall impact of the program through a multi-method assessment design.

Module 1 (Personal Accountability) outcomes are captured through self-efficacy scales in the endline survey, which measure participants' beliefs in their ability to set and achieve goals, manage time effectively, and take ownership of their professional development. We also assess growth mindset using established psychological measures that capture attitudes toward effort and learning from failure.

Module 2 (Professional Communication) outcomes are assessed through a live Zoom-based team challenge where participants must collaborate with peers to solve an engineering design problem and present the solution. Expert evaluators, blind to treatment status, rate participants on their clarity of expression, ability to articulate ideas under time pressure, and effectiveness in professional interactions. Additionally, the endline will include a mock interview, where candidates will participate in a mock-recruitment session with industry professionals.

Module 3 (Technical Expositions and Discussions) outcomes are captured through a coding assessment that requires participants to debug AI-generated code and write original functions. The assessment evaluates not only technical accuracy but also participants' ability to organize their approach, identify key issues, and articulate their reasoning through a structured reflection component in the assessment.

Module 4 (Critical Thinking) outcomes are measured through the problem-solving components of the coding assessment, which require participants to apply logical reasoning to debug errors and develop algorithmic solutions.

Module 5 (Interview Readiness) outcomes are assessed through the endline survey's career confidence measures and mock interview session. These capture participants' preparedness for job interviews, comfort with professional self-presentation, and engagement in career-building activities such as LinkedIn networking.

Beyond the soft skills modules, the technical courses (Python Programming with Git, Building Web Applications, and Data Structures and Algorithms) will be directly assessed through the 45-minute coding assessment in the endline. Finally, the team challenge component of the endline provides a holistic measure of skills that cut across multiple modules, including collaboration, communication under pressure, and the ability to integrate technical and interpersonal competencies in a realistic professional scenario. By examining the effectiveness of individual modules in enhancing skills such as personal accountability, professional communication, technical proficiency, critical thinking, and interview readiness, we can gain insights into the relative importance of each component in the overall training program and subsequent outcomes.

Additionally, we can explore how these skills translate into improved performance in real-world scenarios, such as job interviews and professional interactions. In the present study, however, we are limited by the labor market data of previous cohorts and the baseline data for the latest cohort, but this will help us determine students' confidence level before they undergo the program. This limitation underscores the importance of research to assess the full impact of the training program on participants' confidence, skill development, and labor market outcomes. Nonetheless, our current analysis of the available data will provide insights into patterns of selection into the program, the beliefs and background of participants, and lay the groundwork for our ongoing impact evaluation.

3.3 Data

The data collected in this study includes demographic information from the previous five cohorts of the Program. This includes the number of applications received, the distribution of applicants across zip codes and states, the representation of colleges, the percentage of students from rural areas, and the proportion of families with annual incomes below US$7000 or less than Rs 6 lacs. In addition, we collected data on the selection, enrollment, and labor market outcomes of participants from these cohorts. This includes the number of applicants, the number of selected participants, acceptance rates, and the number of students hired by top engineering firms and the wider industry after graduation from engineering programs. We also collected data on the type of labor market outcomes, such as the percentage of students hired as full-time employees and/or for internships and mentorship opportunities.

In February 2024, prior to the program's start of the sixth cohort, we administered an online survey to gather a wide range of demographic information, including household income, parental education levels, occupations, and geographical distribution based on zip codes. Additionally, we collected academic indicators such as students' standings in engineering colleges, estimates of average and maximum reported salaries of recent graduates, and their performance in entrance exams used for admission criteria. Moreover, the survey assessed psychological traits such as self-efficacy, confidence levels, IQ, EQ scores, and the Big Five Personality traits.

The survey also explored the' career aspirations and professional goals of the participants, identifying perceived barriers to career success, and providing information on the challenges women face in their professional journeys. Furthermore, it captured participants' attitudes toward teamwork and collaboration in academic contexts, along with their participation in group activities and collaboration with individuals of different genders, providing insight into gender dynamics in collaborative environments.

Although this initial survey captured current indicators, subsequent longitudinal surveys will be conducted to track changes over time. We aim to conduct a longitudinal study that spans 2025 to 2028 to track the latest cohort of students.

The survey was complemented with additional data on coding abilities, technical skills, and English proficiency, crucial factors considered in the selection of students, and supplied by the organization. This foundational survey lays the foundation for understanding our participant population and facilitates a comprehensive analysis of the factors that influence their academic and professional trajectories.

In addition to survey-based measures, we use interviewer ratings of candidates' confidence collected during the baseline interviews. These interviews were conducted as part of the program's selection process and took place before randomization. After completion of the baseline surveys and skill assessments, participants (N = 450) were interviewed online through Zoom. Each candidate was interviewed and rated by a single interviewer.

The interviewers assessed the confidence of the candidates using a five-item scale. Each item was scored on a five-point Likert scale ranging from 1 ("not confident at all") to 5 ("extremely confident"). The items capture confidence in non-verbal communication, discussion of technical knowledge, handling challenging questions, presentation of achievements and experiences, and overall communication style. These interviewer-rated measures are conceptually distinct from self-reported confidence collected through surveys. Rather than capturing internal beliefs, the scale reflects how candidates are perceived by external evaluators in a high-stakes selection setting. The psychometric properties of this interviewer-rated confidence scale are examined in the Chapter Appendix.

3.4 Main Results

3.4.1 Representativeness of Applicants in the Previous Five Cohorts

Table 3.3 provides information about the applicant pools across five cohorts. It illustrates the reach of the program in terms of applicant demographics and geographic distribution. Across the five cohorts, the number of applications increased steadily, indicating a growing interest and awareness of the program. This growth is also reflected in the increasing number of zip codes represented, highlighting the program's ability to attract applicants from various regions across India. Applicants include women engineers from all 28 states in India. In addition, the number of colleges represented has increased in each application cycle. For the latest cohort for which we have data, the program received applications from 2,700 different colleges. This represents more than 25% of all colleges in India.[2] This suggests that the program, in its application stage, is able to reach students from a diverse range of educational institutions and regions within India.

The data also shed light on the rural-urban distribution of students, with a slight majority of applicants from rural areas. This indicates the program's reach in remote areas of India. Additionally, a high percentage of applicants come from families with an annual income below US $7,000, which underscores the effectiveness of the program's targeting of economically disadvantaged students. Overall, the data underscore that the program is able to reach students from diverse geographic, educational, and income backgrounds.

3.4.2 Selectivity and Placement of Previous Five Cohorts

Table 3.4 presents data on the selectivity of the program through acceptance rates across different cohorts, with acceptance rates ranging from 0.8% to 1.4%. The program is thus highly selective. For comparison, the acceptance rate at Harvard College, one of the most selective colleges in the United States, is 4%.[3] Similarly, the acceptance rate for admission into the Indian

[2]There are approximately 9,000 colleges in India that provide an engineering or technical degree; For details, see the All India Survey of Higher Education (AISHE) report, 2021-2022 at www.education.gov.in.

[3]For information on the latest cohort of Harvard College: https://college.harvard.edu/admissions/admissions-statistics

Table 3.3: *Applicants Demographic Information From The Previous Five Cohorts of the Program*

Cohort	Applications	Zip Codes	Number of Colleges Represented	Number of States Represented	% Students in Rural Areas	Family Annual Income (Below US$7,000)
2019-2021	7,000	2,100	600	28	50%	83%
2020-2022	15,000	4,600	1,300	28	55%	60%
2021-2023	20,000	5,800	2,800	28	54%	61%
2022-2024	28,000	6,300	2,600	28	57%	83%
2023-2025	22,000	5,900	2,700	28	55%	85%

Notes: This table captures the information of the previous five cohorts of the women engineers training program. The program's inaugural cohort started in 2019.

Institutes of Technology (IITs), India's premier engineering colleges, is 1.3%.[4] This makes the program that we are studying in this paper one of the most competitive of its kind anywhere in the world.

The table also provides information on labor outcomes for participants post-graduation. Specifically, it outlines the number of participants hired by prominent companies, including the "Big-4" (referring to leading tech companies like Google, Facebook, Amazon, and Microsoft), as well as the broader industry. This data suggests the program is effective in preparing participants for successful transitions into the workforce, with a significant percentage securing full-time employment (FTE) opportunities in both prestigious companies and the wider industry. However, given this high level of selectivity, it is difficult to disentangle the source of the program's effectiveness. Is it because of a holistic curriculum taught by experienced faculty in an only-women environment? Or is it simply that those admitted would have strong labor market outcomes regardless of the program? The program's rigorous industry-aligned curriculum, high selectivity, and positive labor market outcomes have motivated the organization that implements the program to conduct an independent impact evaluation. We will cover the design of this evaluation and the baseline survey in more detail in the next section.

[4]More than 1.3 million appear for the IIT entrance exam, of whom 18,000 are admitted. For more, see: https://www.indiatoday.in/education-today/news/story/jee-main-2025-registration-numbers-surge-to-all-time-high-with-13-lakh-applicants-2639594-2024-11-25

Table 3.4: *Applicants and Student Selection, Enrollment and Labor Outcome Information From The Previous Cohorts of WE Program*

Cohort	# Applicants	# Selected Participants	Acceptance %	Hired by Big-4	Hired by Industry
2019-2021	7,000	96	1.4%	9 FTEs (9%)	96 FTEs (100%)
2020-2022	15,000	123	0.8%	28 FTEs (23%)	98 FTEs (80%)
2021-2023	20,000	259	1.3%	33 FTEs (13%)	190 FTEs (75%)
2022-2024	28,000	251	0.9%	17 internship	Data Unavailable

Notes: This table reports acceptance rate in the WE program and the number of students who got recruited by Big-4 (leading tech companies like Google, Facebook, Amazon, and Microsoft) and the wider industry after graduation from engineering programs. FTEs stand for Full-time Employees

3.4.3 Characteristics of Selected Participants

While Table 3.3 presents the demographic composition of applicants, it is important to examine how the characteristics of selected participants compare. Table 3.5 provides information on the socioeconomic background, geographic representation, and institutional origin of students who were admitted to the program across the five cohorts.

The data reveal several noteworthy patterns. First, while the proportion of low-income students among selected participants (28–56%) is lower than among applicants (60–85%), the absolute representation remains substantial for a highly selective program. Given that acceptance rates range from 0.8% to 1.4%, the fact that over one-quarter to one-half of selected participants come from families earning below Rs. 6 lacs annually suggests the program maintains meaningful access for economically disadvantaged students.

Second, although selected participants come from fewer states than the applicant pool (15–22 states compared to all 28), the geographic reach has expanded over time. Recent cohorts draw participants from more than 20 states, indicating broad national coverage despite the program's selectivity.

Perhaps most notably, a substantial majority of selected participants—ranging from 53% to 77%—come from non-elite institutions, defined as colleges outside the IITs, NITs, and NIRF Top 100 engineering rankings.[5] This finding is particularly significant because it suggests the program

[5]The National Institutional Ranking Framework (NIRF) is a framework developed by India's Ministry of Education (MoE) to rank institutions across the country.

provides access to high-quality technical training, mentorship, and professional networks for women engineers who might otherwise lack such opportunities. Students from non-elite colleges might face disadvantages in accessing industry connections and career advancement resources that are more readily available to their peers at premier institutions. By selecting a large proportion of participants from these backgrounds, the program may serve as an important pathway for talented women engineers to overcome institutional and networking barriers and compete for positions at leading technology companies.

Table 3.5: *Characteristics of Selected Participants Across Five Cohorts*

Cohort	Selected Participants	% Low Income	Number of States	% Non-Elite Colleges
2019-2021	96	56.2%	15	77.1%
2020-2022	123	29.3%	19	67.5%
2021-2023	259	52.9%	22	70.7%
2022-2024	251	28.3%	22	55.0%
2023-2025	222	41.9%	22	53.2%

Notes: This table reports characteristics of students selected into the program. Low income is defined as family annual income below Rs. 6 lacs (approximately US$7,000). Non-elite colleges are institutions outside the Indian Institutes of Technology (IITs), National Institutes of Technology (NITs), and NIRF Top 100 engineering rankings. Number of states excludes non-Indian territories.

In conclusion, this training program for women engineers demonstrates diversity and accessibility, both in attracting applicants from varied backgrounds and in selecting participants from across India. Selected students achieve favorable employment outcomes, securing positions in prestigious companies and the broader industry. However, more work is needed to disentangle the impact of the program from selection.

3.4.4 Survey Results of Women in the Sixth Cohort

As noted earlier, we embarked on an evaluation of the program in early 2024. This evaluation involves the sixth cohort (2024-2026). The program is highly oversubscribed, and the number of students who reach the final stages of the selection process are double the program's intake. Thus, we randomly assigned 450 applicants from this final round to treatment and control groups: 225 were selected into the program (treatment), and 225 were not selected (control).

Table 3.6 presents the baseline characteristics of students in the treatment and control groups.

None of the differences between groups are statistically significant. The cohort is predominantly urban (73% live in cities), with about one-third (33%) from families with annual income below Rs. 6 lacs. Most students study in non-elite colleges (68-69%). Compared to the previous cohort (2023-2025, shown in Table 3.5), this cohort has a smaller proportion of students from low-income families but a higher proportion from non-elite colleges. Approximately 8-10% are first-generation engineering students. The cohort also demonstrates strong academic performance: mean scores on the 10th and 12th grade board exams are 94% and 91%, respectively. Standardized measures of coding ability, confidence, IQ, and EQ are balanced across treatment and control groups.[6]

Table 3.6: *Balance Table: Baseline Characteristics by Treatment Status*

	Control Mean (SD)	Treatment Mean (SD)	Difference (p-value)
Urban	0.73 (0.45)	0.72 (0.45)	-0.01 (0.82)
Low Income (<Rs 6 lacs)	0.34 (0.47)	0.32 (0.47)	-0.02 (0.68)
Non-Elite College	0.69 (0.46)	0.68 (0.47)	-0.01 (0.84)
First-Gen Engineering	0.09 (0.29)	0.08 (0.27)	-0.01 (0.72)
10th Grade Score (%)	94.1 (4.2)	93.8 (4.5)	-0.3 (0.45)
12th Grade Score (%)	91.2 (5.8)	91.5 (5.6)	0.3 (0.58)
Coding Ability (std)	0.02 (1.01)	-0.02 (0.99)	-0.04 (0.67)
Confidence (std)	0.01 (0.98)	-0.01 (1.02)	-0.02 (0.83)
IQ Score (std)	-0.03 (1.02)	0.03 (0.98)	0.06 (0.52)
EQ Score (std)	0.02 (0.99)	-0.02 (1.01)	-0.04 (0.66)
N	225	225	

Notes: This table presents baseline characteristics for women randomly assigned to treatment (selected into the program) and control (not selected) groups. Column (1) reports control group means with standard deviations in brackets. Column (2) reports treatment group means with standard deviations in brackets. Column (3) reports the difference between treatment and control means with p-values from two-sided t-tests in parentheses. Low income is defined as annual family income below Rs. 6 lacs (approximately US$7,000). Elite colleges include Indian Institutes of Technology (IITs), National Institutes of Technology (NITs), and institutions ranked in the NIRF Top 100 for engineering. Coding ability and confidence are standardized indices with mean of zero and a standard deviation of one. IQ is measured using Raven's Advanced Progressive Matrices; EQ is measured using Reading the Mind in the Eyes test. Demographics, academic performance, and coding ability are from administrative data (N=225 per group). IQ, EQ, and self-reported confidence are from the baseline survey; statistics for these variables are based on survey respondents only (∼80% of each group). None of the differences are statistically significant at conventional levels, confirming successful randomization.

Beyond these baseline characteristics, Table 3.7 provides additional context on family back-

[6]Technical Score, Aptitude Score, and Coding Ability are measured through assessments that participants took as part of the selection process for the program. Confidence, IQ, and EQ scores were assessed as part of the baseline survey administered by researchers.

ground and perceived barriers to work. Many women in the cohort have mothers who are homemakers (53%), despite the fact that a large proportion of these mothers hold Bachelor's or Master's degrees (61%). The women also identify various obstacles that hinder their path. The barriers most commonly cited include concerns about flexible work schedules (45%), the absence of female mentors (38%), safety concerns (25%), lack of transportation (23%) and gender discrimination (14%). These findings underscore the multifaceted challenges women perceive they face in entering and thriving in engineering fields.

Table 3.7: *Family Background and Perceived Barriers to Work*

Panel A: Family Background	%
Mother is homemaker	53%
Mother has Bachelor's or Master's degree	61%
Father has Bachelor's or Master's degree	72%
Panel B: Perceived Barriers to Work	%
Flexible work schedules	45%
Absence of female mentors	38%
Safety concerns	25%
Lack of transportation	23%
Gender discrimination	14%

Notes: The table presents family background characteristics and perceived barriers to work among women in the sixth cohort. Percentages are calculated among treatment group participants who completed the baseline survey (N=185).

We conduct additional analysis to understand the relationships between cognitive ability, socioeconomic background, and confidence measures. Figure 3.1 presents correlations among five key variables: IQ (measured by Raven's Progressive Matrices), socioeconomic background indicators (low income and first-generation status), and two distinct measures of confidence (interviewer-rated and self-reported). Several patterns emerge from this analysis. First, socioeconomic background is negatively correlated with interviewer-rated confidence: low-income students (r = -0.13) and first-generation students (r = -0.13) are perceived as less confident by interviewers. However, background is essentially uncorrelated with self-reported confidence (r = 0.00 for low income, r = 0.07 for first-generation). Second, interviewer-rated and self-reported confidence are nearly uncorrelated (r = 0.01), confirming that these measures capture distinct constructs—external perceptions versus internal beliefs. Third, IQ shows no meaningful correlation with either background variables or confidence measures, suggesting that cognitive ability operates independently of socioeconomic status and confidence in this sample. In the next section, I further examine this gap in confidence that disadvantages students from lower socioeconomic backgrounds.

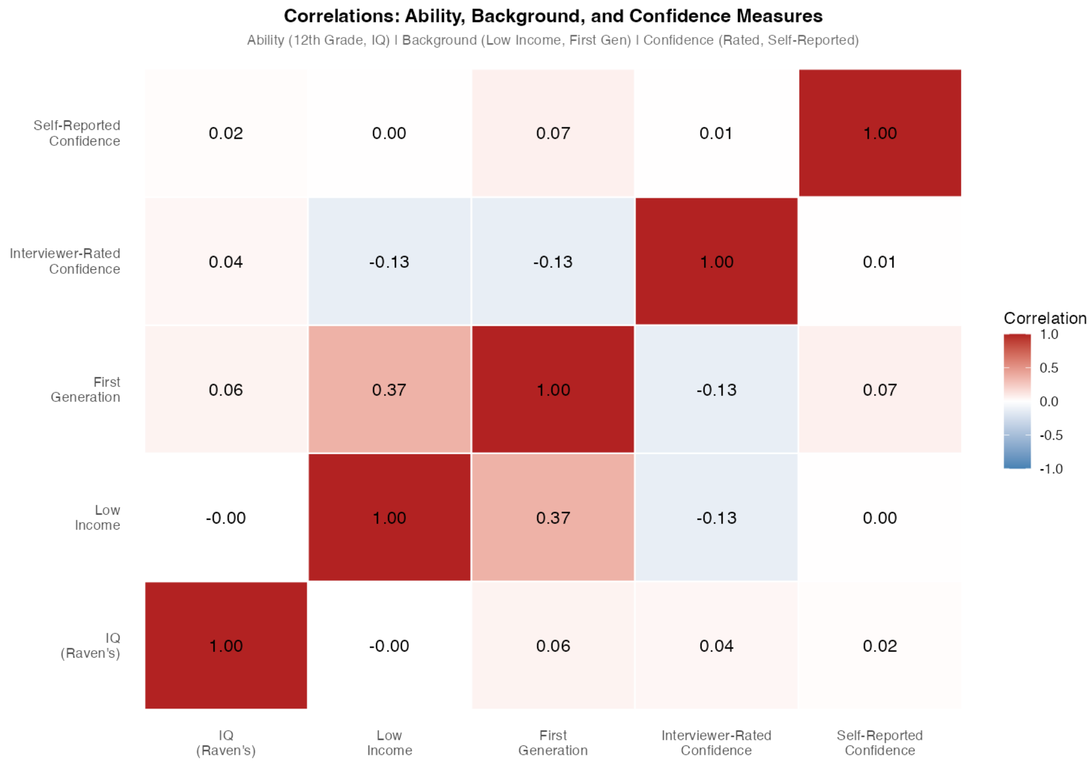


Figure 3.1: *Correlations Among Cognitive Ability, Background, and Confidence Measures*

Notes: This figure presents pairwise correlations among cognitive ability (IQ measured by Raven's Progressive Matrices), socioeconomic background (low income defined as family income below Rs. 6 lacs; first-generation defined as neither parent having a college degree), and confidence measures (interviewer-rated confidence from baseline interviews; self-reported confidence from the General Self-Efficacy Scale). The sample includes the 360 participants who completed the baseline survey.

3.4.5 Interviewer-Rated Confidence

In addition to survey-based measures of self-reported confidence, we collected interviewer ratings of candidates' confidence during the baseline interviews. Interviews are an important part of selection processes across industries and countries. Interviews often determine whether an individual gets a scholarship, a coveted job, or admission into a selective educational program. Research has shown that evaluators' implicit biases can affect their assessments of candidates

(Carlana, 2019). One factor that might determine success or failure in an interview is how a candidate's confidence is perceived by the interviewer(s). An interviewer's perception of a candidate's confidence might influence their decision regarding whether to hire, admit, or offer a scholarship.

Women, in particular, might be vulnerable to the penalty of being perceived as less confident (Alan et al., 2018). Recent research suggests a gender-based confidence gap among women in labor market settings (Exley and Nielsen, 2024; Coffman et al., 2023, 2024). Perceptions and external ratings of women's confidence might affect their opportunities in the labor market and more generally exacerbate gender-based disparities in male-dominated STEM fields.

After completion of baseline surveys and skill assessments, participants (N = 450) were interviewed online through Zoom as part of the program's selection process. Each candidate was interviewed and rated by a single interviewer. In total, seventeen interviewers rated candidates, with the number of candidates per interviewer ranging from 1 to 108. Interviewers rated participants' confidence on a 1-5 scale (1 = not confident at all, 5 = extremely confident) across five items: (a) ability to project confidence through non-verbal cues such as body language and eye contact; (b) confidence in discussing technical knowledge; (c) confidence in handling challenging questions; (d) confidence in presenting achievements and experiences; and (e) overall communication style.

Table 3.8 presents summary statistics for each item. On average, participants scored highest on non-verbal confidence (mean = 4.00) and lowest on handling challenging questions (mean = 3.58). The overall mean across items is 3.80, indicating that most candidates were perceived as "Confident" or higher by interviewers.

Table 3.8: *Interviewer-Rated Confidence: Summary Statistics*

Item	Mean	SD
Non-verbal communication	4.00	0.84
Technical knowledge	3.70	0.97
Handling challenging questions	3.58	1.05
Presenting achievements	3.81	0.88
Communication style	3.88	0.90
Overall (average across items)	3.80	0.80

Notes: This table presents summary statistics for the five-item interviewer-rated confidence scale. Each item was scored on a 5-point scale (1 = not confident at all, 5 = extremely confident). N = 450.

Figure 3.2 presents the distribution of rated confidence scores by family income. We find that rated confidence is negatively correlated with low-income status (r = -0.13, $p < 0.01$) and first-generation student status (r = -0.13, $p < 0.05$). Students from low-income families have a mean rated confidence score of 3.64 compared to 3.87 for students from higher-income families. Similarly, first-generation students have a mean score of 3.66 compared to 3.85 for non-first-generation students. While the relationship of rated confidence with these covariates is modest, the directionality suggests that the instrument is picking up advantages that family background and social capital can bestow on how individuals' confidence is perceived in interview settings.

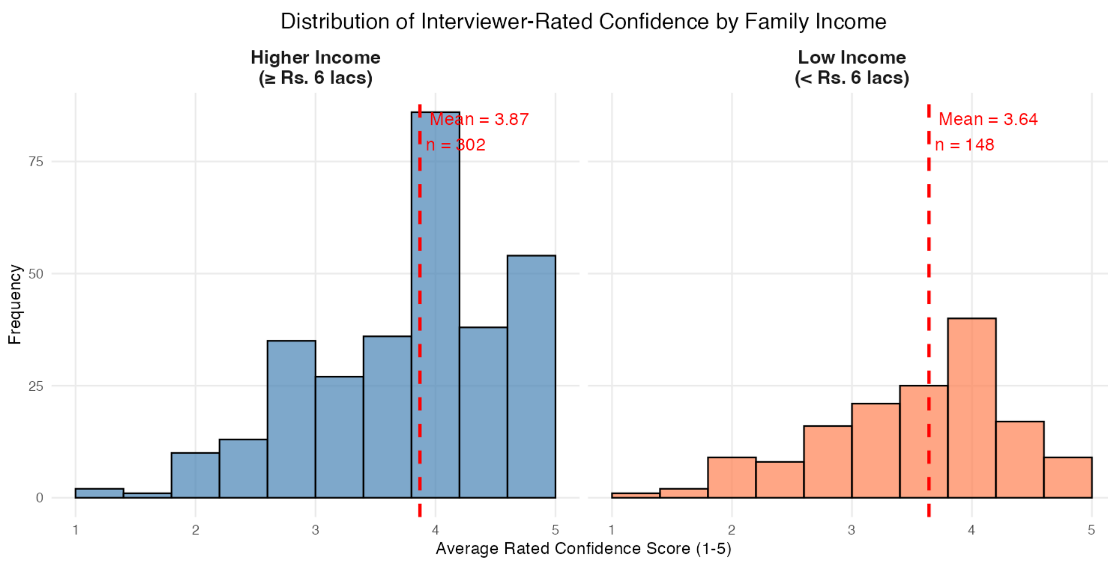


Figure 3.2: *Distribution of Interviewer-Rated Confidence by Family Income*

Notes: This figure presents the distribution of average interviewer-rated confidence scores by family income level. Low income is defined as annual family income below Rs. 6 lacs (approximately US$7,000). The difference in means between groups is statistically significant ($p < 0.01$). N = 450.

These interviewer-rated measures are conceptually distinct from self-reported confidence collected through surveys. Rather than capturing internal beliefs, the scale reflects how candidates are perceived by external evaluators in a high-stakes selection setting. We find that the correlation between interviewer-rated confidence and self-reported confidence (measured using the General Self-Efficacy Scale) is essentially zero (r = 0.005, p = 0.92), confirming that these measures capture distinct constructs.

This distinction has important implications when we examine confidence by socioeconomic background. Figure 3.3 presents a comparison of self-reported and interviewer-rated confidence by

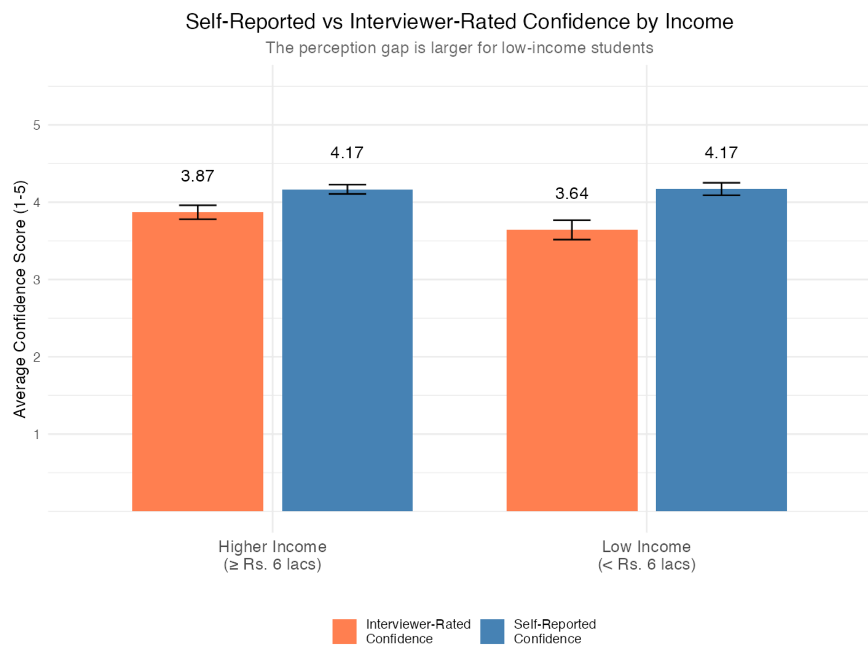


Figure 3.3: *Self-Reported vs Interviewer-Rated Confidence by Family Income*

Notes: This figure compares average self-reported confidence (from the General Self-Efficacy Scale) and interviewer-rated confidence by family income level. Error bars represent 95% confidence intervals. While self-reported confidence is identical across income groups ($p = 0.95$), interviewer-rated confidence is significantly lower for low-income students ($p < 0.01$). The difference in the "perception gap" (interviewer-rated minus self-reported) between income groups is statistically significant ($p = 0.03$). N = 360.

family income level. I find that while self-reported confidence is virtually identical across income groups (mean = 4.17 for both low-income and higher-income students, $p = 0.95$), interviewer-rated confidence differs significantly by income (3.64 for low-income vs. 3.87 for higher-income students, $p < 0.01$). This creates what we term a "perception gap"—the difference between how confident students feel internally and how confident they are perceived by external evaluators.

The perception gap is significantly larger for low-income students (-0.51) compared to higher-income students (-0.28), a difference that is statistically significant ($p = 0.03$). In other words, while all students tend to rate themselves as more confident than interviewers perceive them to be, this discrepancy is nearly twice as large for students from economically disadvantaged backgrounds. This finding suggests that low-income students may face a "perception penalty" in interview settings: despite having the same internal confidence as their higher-income peers, they are perceived as less confident by external evaluators. Such perceptions could disadvantage low-income women in competitive selection processes for scholarships, jobs, and educational programs.

We intend to use the interviewer-rated confidence instrument as an outcome variable in the endline to evaluate whether participation in the program increases how confident women participants are perceived by external evaluators, and whether the program can reduce the perception gap faced by students from low-income backgrounds.

3.5 Conclusion

This paper examines a unique all-women STEM program in India that aims to reduce barriers for women in STEM. Using data from six cohorts of the program, we find that it successfully targets women engineering students from a wide range of geographical and socioeconomic backgrounds. For instance, a substantial minority of the program participants are from rural and low-income backgrounds, and a majority of them are from India's non-elite colleges. The program also receives thousands of applications for a few hundred slots, making it one of the most selective programs of its kind.

The survey results of women in the sixth cohort highlight demographic trends and perceived obstacles to career advancement. Despite being predominantly urban, women participants report facing challenges such as the absence of female mentors and mobility constraints. This strengthens the case that this mostly online program, which provides access to mentoring, might reduce certain barriers for women.

We also use interview data from the baseline survey of the sixth cohort to develop and analyze a novel measure of confidence. This measure captures confidence as perceived by external evaluators. Using these data, we find a "perception gap" in confidence across income groups. While self-reported confidence is identical across income groups, interviewers perceive low-income students as significantly less confident than their higher-income peers. This perception gap is nearly twice as large for low-income students, suggesting that socioeconomic background influences how external evaluators perceive candidates in high-stakes interview settings—even when those candidates have the same internal confidence. This finding suggests that women from a low-income background might face additional barriers to success in high-stakes STEM settings.

Chapter 3 Appendix: Psychometric Properties of Interviewer-Rated Confidence Scale

The interviewer-rated confidence scale demonstrates good psychometric properties. The five items show high internal consistency (Cronbach's $\alpha = 0.91$), suggesting they measure a coherent underlying construct. Factor Analysis of the instrument indicates that the instrument is comprised of a single factor. The first factor accounts for close to 95% of the total variation.

In total, there are 17 interviewers. Each interviewer was assigned a candidate to interview and also to rate them on the confidence instrument. The number of candidates rated by an interviewer ranged from 1 to 108. To understand the within and between interviewer variance on scores, I ran two random effects models, one with sum scores and another with IRT theta scores as the outcome. Before this analysis, I dropped the two interviewers who only interviewed one candidate each. Based on these results, I calculate the Intracluster Correlation Coefficient (ICC) for each model. For the random effects model with theta scores as the outcome, the ICC is 0.003. Therefore, only 0.3% of the variation in theta scores is between (vs. within) interviewers. The ICC is similarly low (ICC = 0.006) for the random effects model with sum scores as the outcome. This shows a high degree of consistency in ratings between interviewers.

Appendix A

Data Sources and Methodology

A.1 Chapter 1: How Leaders Emerge

Data Collection Period: May-July 2025

Sample: 203 mixed-gender teams (610 students) at an engineering college in rural India

IRB Approval: Harvard IRB approval IRB25-0406

Pre-registration: AEA RCT Registry (AEARCTR-0016103)

Primary Data Sources:

- Baseline survey of individual characteristics
- PAGE (Perceiving AI-Generated Emotions) assessment
- Endline survey of team dynamics and leadership
- Project evaluation scores from blind expert judges

A.2 Chapter 2: Misperceptions about Caste

Data Collection Period: August 2023

Sample: 774 college-aged respondents at a computer skills institute in Central India

IRB Approval: Harvard IRB approval IRB23-0502

Pre-registration: AEA RCT Registry (AEARCTR-0011706)

Primary Data Sources:

- Qualtrics survey measuring perceptions about caste
- Randomized information treatment and placebo
- Endline survey measuring attitudes and preferences

External Data Sources:

- Pew Research Center survey on caste in India (2021) for population estimates
- National Sample Survey (NSS) for income and education statistics

A.3 Chapter 3: Women in STEM

Data Collection Period: 2019-2024 (administrative data); February 2024 (baseline survey)

Sample:

- Administrative data: Five cohorts (2019-2025) with 92,000+ applicants
- Baseline survey: 450 participants in sixth cohort (2024-2026)

IRB Approval: Harvard IRB approval IRB25-1425 **Pre-registration:** Ongoing RCT registered with AEA RCT Registry

Primary Data Sources:

- Program administrative records (applications, selections, placements)
- Baseline survey of demographic, academic, and psychological characteristics
- Interviewer-rated confidence assessments
- Cognitive assessments (Raven's Progressive Matrices, Reading the Mind in the Eyes)

Appendix B

Survey Instruments

Complete survey instruments for each chapter are available upon request from the author. Key measures used across chapters include:

Psychological Measures:

- General Self-Efficacy Scale (Schwarzer & Jerusalem, 1995)
- Big Five Personality Inventory (short form)
- Growth Mindset Scale (Dweck, 2006)

Cognitive Measures:

- Raven's Advanced Progressive Matrices (IQ)
- Reading the Mind in the Eyes Test (EQ)
- PAGE - Perceiving AI-Generated Emotions (Chapter 1)

Attitudes and Preferences:

- Attitudes toward caste index (Chapter 2)
- Preferences for affirmative action index (Chapter 2)
- Team equity and effectiveness indices (Chapter 1)
- Leadership effectiveness index (Chapter 1)

References

Acemoglu, D., D. Laibson, and J. A. List (2014, May). Equalizing Superstars: The Internet and the Democratization of Education. *American Economic Review 104*(5), 523–527.

Ahmed, H., M. Mahmud, F. Said, and Z. Tirmazee (2024, January). Encouraging Female Graduates to Enter the Labor Force: Evidence from a Role Model Intervention in Pakistan. *Economic Development and Cultural Change 72*(2), 919–957. Publisher: The University of Chicago Press.

Alan, S., S. Ertac, and I. Mumcu (2018, December). Gender Stereotypes in the Classroom and Effects on Achievement. *The Review of Economics and Statistics 100*(5), 876–890.

Alesina, A., A. Miano, and S. Stantcheva (2023, January). Immigration and Redistribution. *The Review of Economic Studies 90*(1), 1–39.

Alesina, A. and S. Stantcheva (2020). Diversity, immigration, and redistribution. *AEA Papers and Proceedings 110*, 329–334.

Alesina, A., S. Stantcheva, and E. Teso (2018, February). Intergenerational Mobility and Preferences for Redistribution. *American Economic Review 108*(2), 521–554.

Aparicio Fenoll, A. and S. Zaccagni (2022, December). Gender mix and team performance: Differences between exogenously and endogenously formed teams. *Labour Economics 79*, 102269.

Apesteguia, J., G. Azmat, and N. Iriberri (2012a, January). The Impact of Gender Composition on Team Performance and Decision Making: Evidence from the Field. *Management Science 58*(1), 78–93. Publisher: INFORMS.

Apesteguia, J., G. Azmat, and N. Iriberri (2012b, January). The Impact of Gender Composition on Team Performance and Decision Making: Evidence from the Field. *Management Science 58*(1), 78–93. Publisher: INFORMS.

Ashraf, N., N. Bau, C. Low, and K. McGinn (2020, May). Negotiating a Better Future: How Interpersonal Skills Facilitate Intergenerational Investment*. *The Quarterly Journal of Economics 135*(2), 1095–1151.

Bagde, S., D. Epple, and L. Taylor (2016, June). Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India. *American Economic Review 106*(6), 1495–1521.

Basiglio, S., D. Del Boca, and C. Pronzato (2024, June). The Impact of the 'Coding Girls' Program on High School Students' Skills, Awareness and Aspirations. *CESifo Economic Studies*, ifae006.

Berge, L. I. O., K. S. Juniwaty, and L. H. Sekei (2016, November). Gender composition and group dynamics: Evidence from a laboratory experiment with microfinance clients. *Journal of Economic Behavior & Organization 131*, 1–20.

Bertrand, M., R. Hanna, and S. Mullainathan (2010, February). Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics 94*(1), 16–29.

Booth, A. L., L. Cardona-Sosa, and P. Nolen (2018, December). Do single-sex classes affect academic achievement? An experiment in a coeducational university. *Journal of Public Economics 168*, 109–126.

Born, A., E. Ranehill, and A. Sandberg (2022, March). Gender and Willingness to Lead: Does the Gender Composition of Teams Matter? *The Review of Economics and Statistics 104*(2), 259–275.

Bostwick, V. K. and B. A. Weinberg (2022, April). Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs. *Journal of Labor Economics 40*(2), 397–436. Publisher: The University of Chicago Press.

Brandts, J., D. J. Cooper, and R. A. Weber (2015, November). Legitimacy, Communication, and Leadership in the Turnaround Game. *Management Science 61*(11), 2627–2645. Publisher: INFORMS.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2023, July). How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools. *The Economic Journal 133*(653), 1773–1809.

Béteille, A. (2012). The Peculiar Tenacity of Caste. *Economic and Political Weekly 47*(13), 41–48. Publisher: Economic and Political Weekly.

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

Carlana, M. and M. Fort (2022, May). Hacking Gender Stereotypes: Girls' Participation in Coding Clubs. *AEA Papers and Proceedings 112*, 583–587.

Castillo, M., R. Petrie, M. Torero, and L. Vesterlund (2013, March). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics 99*, 35–48.

Ceci, S. J. and W. M. Williams (2011, February). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences 108*(8), 3157–3162. Publisher: Proceedings of the National Academy of Sciences.

Chemin, M. (2021, June). Does appointing team leaders and shaping leadership styles increase effort? Evidence from a field experiment. *Journal of Economic Behavior & Organization 186*, 12–32.

Chen, J. and D. Houser (2019, December). When are women willing to lead? The effect of team gender composition and gendered tasks. *The Leadership Quarterly 30*(6), 101340.

Coffman, K., M. P. Ugalde Araya, and B. Zafar (2024). A (dynamic) investigation of stereotypes, belief-updating, and behavior. *Economic Inquiry 62*(3), 957–983. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecin.13219.

Coffman, K. B., C. L. Exley, and M. Niederle (2023). The role of beliefs in driving gender discrimination. *Management Science 69*(5), 2681–2700.

De Paola, M., F. Gioia, and V. Scoppa (2022, September). Female leadership: Effectiveness and perception. *Journal of Economic Behavior & Organization 201*, 134–162.

Deming, D. J. (2017, November). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Desai, S. and A. Dubey (2012, March). Caste in 21st Century India: Competing Narratives. *Economic and political weekly 46*(11), 40–49.

Deserranno, E., M. Stryjan, and M. Sulaiman (2019, October). Leader Selection and Service Delivery in Community Groups: Experimental Evidence from Uganda. *American Economic Journal: Applied Economics 11*(4), 240–267.

Deshpande, A. (2016). Double jeopardy? Caste, affirmative action, and stigma. Working Paper 2016/71, Helsinki: The United Nations University World Institute for Development Economics Research (UNU-WIDER). ISBN: 9789292561147.

Deshpande, A. and R. Ramachandran (2019, June). Traditional hierarchies and affirmative action in a globalizing economy: Evidence from India. *World Development 118*, 63–78.

Eccles, J. S. and A. Wigfield (2002, February). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology 53*(Volume 53, 2002), 109–132. Publisher: Annual Reviews.

Eckel, C., L. Gangadharan, P. J. Grossman, and N. Xue (2021, July). Chapter 7: The gender leadership gap: insights from experiments. Section: A Research Agenda for Experimental Economics.

Edin, P.-A., P. Fredriksson, M. Nybom, and B. Öckert (2022, April). The Rising Return to Noncognitive Skill. *American Economic Journal: Applied Economics 14*(2), 78–100.

Eisenkopf, G., Z. Hessami, U. Fischbacher, and H. Ursprung (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization 115*(C), 123–143. Publisher: Elsevier.

Englmaier, F., S. Grimm, D. Grothe, D. Schindler, and S. Schudy (2024, August). The Effect of Incentives in Nonroutine Analytical Team Tasks. *Journal of Political Economy 132*(8), 2695–2747. Publisher: The University of Chicago Press.

Exley, C. L. and K. Nielsen (2024, March). The Gender Gap in Confidence: Expected but Not Accounted For. *American Economic Review 114*(3), 851–885.

Frisancho, V. and K. Krishna (2016, May). Affirmative action in higher education in India: targeting, catch up, and mismatch. *Higher Education 71*(5), 611–649. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 5 Publisher: Springer Netherlands.

Gangadharan, L., T. Jain, P. Maitra, and J. Vecci (2016, November). Social identity and governance: The behavioral response to female leaders. *European Economic Review 90*, 302–325.

Glick, P. and S. T. Fiske (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology 70*(3), 491–512. Place: US Publisher: American Psychological Association.

Grigorieff, A., C. Roth, and D. Ubfal (2020, June). Does Information Change Attitudes Toward Immigrants? *Demography 57*(3), 1117–1143.

Hardt, D., L. Mayer, and J. Rincke (2025, May). Who Does the Talking Here? The Impact of Gender Composition on Team Interactions. *Management Science 71*(5), 4131–4152. Publisher: INFORMS.

Hasan, S. and S. Bagde (2013, December). The Mechanics of Social Capital and Academic Performance in an Indian College. *American Sociological Review 78*(6), 1009–1032. Publisher: SAGE Publications Inc.

Hnatkovska, V., A. Lahiri, and S. B. Paul (2013, March). Breaking the Caste Barrier: Intergenerational Mobility in India. *Journal of Human Resources 48*(2), 435–473. Publisher: University of Wisconsin Press Section: Articles.

Hoogendoorn, S., H. Oosterbeek, and M. van Praag (2013, July). The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment. *Management Science 59*(7), 1514–1528. Publisher: INFORMS.

Hunt, J. (2016, January). Why do Women Leave Science and Engineering? *ILR Review 69*(1), 199–226. Publisher: SAGE Publications Inc.

Jaffrelot, C. (2010). Caste and Politics. *India International Centre Quarterly 37*(2), 94–116. Publisher: India International Centre.

Kahn, S. and D. Ginther (2018, July). Women and Science, Technology, Engineering, and Mathematics (STEM): Are Differences in Education and Careers Due to Stereotypes, Interests, or Family? In S. L. Averett, L. M. Argys, and S. D. Hoffman (Eds.), *The Oxford Handbook of Women and the Economy*, pp. 0. Oxford University Press.

Kane, J. V. and J. Barabas (2019). No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments. *American Journal of Political Science 63*(1), 234–249. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12396.

Kantor, P. (2002, June). Female mobility in India: the influence of seclusion norms on economic outcomes. *International Development Planning Review 24*(2), 145–160. Publisher: Liverpool University Press (UK).

Karpowitz, C. F., S. D. O'Connell, J. Preece, and O. Stoddard (2024, September). Strength in Numbers? Gender Composition, Leadership, and Women's Influence in Teams. *Journal of Political Economy 132*(9), 3077–3114. Publisher: The University of Chicago Press.

Kraus, M. W., J. M. Rucker, and J. A. Richeson (2017, September). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences 114*(39), 10324–10331. Publisher: Proceedings of the National Academy of Sciences.

Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015, April). How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments. *American Economic Review 105*(4), 1478–1508.

Lamiraud, K. and R. Vranceanu (2018, January). Group gender composition and economic decision-making: Evidence from the *Kallystée* business game. *Journal of Economic Behavior & Organization 145*, 294–305.

Lavy, V. and A. Schlosser (2011, April). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics 3*(2), 1–33.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies 76*(3), 1071–1102. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

Levy, D. M., K. Padgitt, S. J. Peart, D. Houser, and E. Xiao (2011, January). Leadership, cheap talk and *really* cheap talk. *Journal of Economic Behavior & Organization 77*(1), 40–52.

Mernyk, J. S., S. L. Pink, J. N. Druckman, and R. Willer (2022, April). Correcting inaccurate metaperceptions reduces Americans' support for partisan violence. *Proceedings of the National Academy of Sciences 119*(16), e2116851119. Publisher: Proceedings of the National Academy of Sciences.

Montes, F., R. C. Jimenez, and J.-P. Onnela (2018, October). Connected but segregated: social networks in rural villages. *Journal of Complex Networks 6*(5), 693–705.

Mouganie, P. and Y. Wang (2020, July). High-Performing Peers and Female STEM Choices in School. *Journal of Labor Economics 38*(3), 805–841. Publisher: The University of Chicago Press.

Oosterbeek, H. and R. van Ewijk (2014, February). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review 38*, 51–63.

Pathania, G. J. and W. G. Tierney (2018, September). An ethnography of caste and class at an Indian university: creating capital. *Tertiary Education and Management 24*(3), 221–231.

Porter, C. and D. Serra (2020, July). Gender Differences in the Choice of Major: The Importance of Female Role Models. *American Economic Journal: Applied Economics 12*(3), 226–254.

Post, D. (2016, February). Affirmative Action Matters: Creating Opportunities for Students around the World edited by Laura Dudley Jenkins and Michele S. Moses. *Comparative Education Review 60*(1), 184–186. Publisher: The University of Chicago Press.

Prakash, N. (2020, December). The impact of employment quotas on the economic lives of disadvantaged minorities in India. *Journal of Economic Behavior & Organization 180*, 494–509.

Reuben, E. and K. Timko (2018, December). On the effectiveness of elected male and female leaders and team coordination. *Journal of the Economic Science Association 4*(2), 123–135.

Siddique, Z. (2011, December). Evidence on Caste Based Discrimination. *Labour Economics 18*, S146–S159.

Subramanian, A. (2015, April). Making Merit: The Indian Institutes of Technology and the Social Life of Caste. *Comparative Studies in Society and History 57*(2), 291–322. Publisher: Cambridge University Press.

Sutter, M., R. Bosman, M. G. Kocher, and F. v. Winden (2009, September). Gender pairing and bargaining—Beware the same sex! *Experimental Economics 12*(3), 318–331.

Sönmez, T. and M. B. Yenmez (2019, March). Affirmative Action in India via Vertical and Horizontal Reservations. *Boston College Working Papers in Economics*. Number: 977 Publisher: Boston College Department of Economics.

Thapar, R. (1999, November). *From Lineage to State: Social Formations in the Mid-First Millennium B.C. in the Ganga Valley*. Oxford University Press. Google-Books-ID: WCwpDwAAQBAJ.

Thorat, S. and P. Attewell (2007). The Legacy of Social Exclusion: A Correspondence Study of Job Discrimination in India. *Economic and Political Weekly 42*(41), 4141–4145. Publisher: Economic and Political Weekly.

Truffa, F. and A. Wong (2025, July). Undergraduate Gender Diversity and the Direction of Scientific Research. *American Economic Review 115*(7), 2414–2448.

Vial, A. C., J. L. Napier, and V. L. Brescoll (2016, June). A bed of thorns: Female leaders and the self-reinforcing cycle of illegitimacy. *The Leadership Quarterly 27*(3), 400–414.

Vijay, D. and V. G. Nair (2022, August). In the Name of Merit: Ethical Violence and Inequality at a Business School. *Journal of Business Ethics 179*(2), 315–337.

Weidmann, B. and D. J. Deming (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica 89*(6), 2637–2657. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA18461.

Weidmann, B., J. Vecci, F. Said, D. J. Deming, and S. R. Bhalotra (2024, July). How Do You Find a Good Manager?

Weidmann, B. and Y. Xu (2024, September). PAGE: A Modern Measure of Emotion Perception for Teamwork and Management Research. arXiv:2410.03704 [cs].

Yang, Y., T. Y. Tian, T. K. Woodruff, B. F. Jones, and B. Uzzi (2022, September). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences 119*(36), e2200841119. Publisher: Proceedings of the National Academy of Sciences.