

C A G E

Tackling Sexual Harassment: Short and Long-Run Experimental Evidence from India

CAGE working paper no. 728

October 2024

Karmini Sharma

Tackling Sexual Harassment: Short and Long-Run Experimental Evidence from India*

Karmini Sharma[†]

Abstract

Sexual harassment awareness training is a key tool to combat sexual harassment, which affects nearly 205 million people in the workplace (ILO, 2022). This paper provides the first randomized evaluation of such training in collaboration with colleges in Delhi, India, to study its impact on sexual harassment. I randomly assigned men to receive this training, with empathy-building, and collected reports of sexual harassment from women in their classes. The training significantly reduced sexual harassment for upto 3 years and altered men's perceptions of social disapproval more than their intrinsic attitudes. It also led to a long-lasting reduction in classroom romantic relationships. A mechanism experiment suggests this is due to women finding it difficult to judge men's quality when social disapproval generates a pooling equilibrium. A similar intervention for women had no detectable effects. Finally, men's training increased women's labor market engagement without affecting their mental well-being or test scores.

JEL codes: D91, J16, J28, K42, O12.

Keywords: Gender, sexual harassment, deterrence, beliefs.

*Funding for this project was provided by the UK Foreign, Commonwealth Development Office, awarded through the Crime and Violence Initiative at J-PAL. I thank J-PAL-SA, IWWAGE-IFMR, and Safecity for support. I am indebted to Pooja, Najma, Rahul, and Nazmeen for field work and Prakriti Singh, Ayushmita Samal and Laura Lufrey for research assistance. I thank Vanshika, Siddharth, and Kajal for excellent training delivery. I thank James Fenske and Clement Imbert, for their guidance on the project and Debraj Ray for advise on the theoretical framework. I thank Zeinab Aboutalebi, Arun Advani, Robert Akerlof, Anik Ashraf, Haseeb Ashraf, Abhijeet Bannerjee, Lori Beaman, Sonia Bhalotra, Apurav Bhatiya, Girija Borker, Nina Buchmann, Stefano Caria, Costas Cavounidis, Pascaline Dupas, Marcel Fafchamps, Olle Folke, Felix Forster, Maitreesh Ghatak, Bishnu Gupta, Seema Jayachandran, Gabriel Kreindler, Anandi Mani, Sharun Mukand, Muriel Niederle, Dev Patel, Gautam Rao, Johanna Rickne, Chris Roth, Andreas Stegmann, Teodora Tsankova, and Todor Tochev for feedback. IRB approval was from IFMR and University of Warwick and pre-registry at AEA registry (AEARCTR-0003678). All errors are my own.

[†]Imperial College London Business School, ksharma3[at]imperial[dot]ac[dot]uk

1 Introduction

Nearly 71% of women face sexual violence and harassment at least once in their lifetime.¹² Sexual harassment and violence can have a massive impact on survivors' psychological, economic, and social lives (Borker, 2017; Talboys et al., 2017; Fitzgerald and Cortina, 2018; Folke and Rickne, 2020; Azmat et al., 2020; Adams-Prassl et al., 2022). For example, victims are more likely to quit their jobs or choose lower-quality educational institutions to avoid sexual harassment (Borker, 2017; Folke and Rickne, 2020; Batut et al., 2021). This can lead to misallocation of talent in an economy, affecting its GDP growth (Hsieh et al., 2019). Since women are more likely to be the victims of sexual harassment, it further contributes to gender gaps in education, labour force participation and earnings.

Despite this, there is little rigorous evidence on what can help reduce sexual harassment within interpersonal environments like workplaces and educational institutions, where around 205 million workers will suffer from sexual harassment at least once (ILO, 2022). Billions of dollars have been spent on sexual harassment awareness training, which is commonly used to prevent such incidents. Yet, to the best of my knowledge, there is no rigorous study on its effectiveness in any context, let alone for the Global South.³ There has been research studying effects of this training on some possible mechanisms, such as changes in self-reported attitudes or awareness in high income countries, but none study the causal impact on sexual harassment incidence (Roehling and Huang (2018); UN (2020)).

In this paper, I provide the first experimental evidence on both short- and long-run effects of sexual harassment awareness training for men on the incidence of sexual harassment reported by women in colleges in Delhi. In 2020–2021, more than 2 million women enrolled in undergraduate courses in India (Ministry of Education, 2020-21), highlighting the importance of studying sexual harassment in this context. I collaborated with three different colleges that together account for 28,203 enrolled students in a university in Delhi. The training was provided by a well-known NGO, Safecity, to men in randomly selected classes in two of the colleges. This initiative primarily focused on men, who are more likely to perpetrate harassment (Guttek et al., 1990).

The training was conducted in person, while many existing training programs at universities are either online, self-paced or non-existent. This in-person approach may

¹Source: <https://asiapacific.unwomen.org/en/media-and-resources/evaw-facts-and-figures>.

²Sexual harassment is defined as any unwanted and unwelcome behavior of a sexual nature (Sexual Harassment of Women at Workplace Act, 2013, India).

³This training is mandatory in India, most of the US states, South Korea, France and is considered the main tool to reduce sexual harassment in majority of the world (Dobbins and Klev, 2020).

be particularly desirable since the treatment of social sanction may play an important role in deterring harassment. When individuals are aware that their peers consider their behavior reprehensible and have been instructed to support victims, the perceived cost of engaging in that behavior becomes strong, even if personal attitudes toward sexual harassment remain unchanged after the training. This aspect may be particularly vital in educational settings like college campuses, where peer pressure is notoriously high.⁴

The training, which lasted up to five hours, had two parts: awareness and empathy-building. The first part included a discussion of sexual harassment laws and ways to support survivors. The second part included a discussion of anonymous narratives of women in men's colleges, detailing the impact sexual harassment has had on them. While the first component is common to most general sexual harassment training, the second component was added to reduce possible backlash effects on attitudes (Bingham and Scherer, 2001; Dobbin and Kalev, 2020).

Studying the impact of sexual harassment training presents several challenges. First, the scarcity of data on sexual harassment prevalence makes it difficult to study its prevention. For instance, there were only 313 cases reported from colleges, while the prevalence is estimated to affect up to 1.84 million women per year.⁵ Second, it is difficult to directly link sexual harassment outcomes of potential victims to the training of potential perpetrators, as regular contact between the two groups is necessary for detecting effects. Third, voluntary sign-up for training can lead to self-selection, potentially resulting in biased estimates.

This study overcomes these challenges. First, women reported sexual harassment via a well-tested questionnaire that reduces victim-backlash and survey fatigue (developed and tested by Fitzgerald, 1988; Fitzgerald et al., 1995).⁶ Second, women were asked to report if someone from their class had harassed them, without revealing the perpetrator's identity, thus eliminating the risk of backlash while still maintaining a tight link between the training and reports of sexual harassment. Additionally, a teacher was quasi-randomly selected to provide their class slot for either training or survey, unknown to students, which led to minimal student dropout (less than 3%) and mitigating self-selection concerns. Crucially, this is the first study to collect and utilize panel data

⁴See Bursztyn and Jensen (2015) for the role of peer pressure, Zhu (2019) for the importance of peers and networks formed in college, and Beaman and Magruder (2012) and Sacerdote (2001) for the utility of such relationships in the labor market.

⁵Number of reports comes from AISHE reports (2020-21) on cases reported and disposed off, and a prevalence rate of 92% from baseline rates of sexual harassment.

⁶New methods have been used in some work such as garbling Boudreau et al. (2023) and more costly third party observers Amaral et al. (2023). Former method would not identify class level sexual harassment and the latter is not possible in a class based setting without triggering demand and Hawthorne effects for perpetrators.

on both male and female students, after the training, on sexual harassment and other outcomes to study the impact of the training.

To ensure consistent understanding of sexual harassment across treatment and control classes, all women were informed about it at baseline. This context, therefore, allowed for assessing the impact of male training when women were already informed. Additionally, a separate female intervention was conducted in another college to evaluate the effects of providing this information to women alone. Follow-up surveys were conducted three to four months and two to three years after the interventions to study short- and long-run effects in all the collaborating colleges.

I begin by estimating the short-run effects. The male intervention results in a significant 0.06 standard deviation (sd) reduction in overall sexual harassment. Specifically, women reported a complete elimination of extreme forms of harassment, decreasing from 1% to 0% incidence, translating to 52 fewer women among 1,200 experiencing extreme sexual harassment over an academic year. Notably, extreme forms of sexual harassment include groping, pinching, and other extreme acts without the victim's consent and generate very high costs for victims. Conversely, the female intervention showed no significant effect on harassment.

This setting allows for the study of the training's impact on other forms of interactions between men and women that could affect future socio-economic outcomes (Zhu, 2019). Hence, I then analyzed how the training impacted different types of relationships, including romantic and platonic interactions between men and women. Participants also engaged in lab-in-the-field games, to decide to partner with either same- or opposite-sex participants. The training led to a significant 0.13 sd decrease in opposite-sex relationships, particularly within-classroom romantic relationships, which dropped by 1.3 percentage points from a baseline of 2.3% in the control group.

Using a signalling game for interactions between men and women, and incorporating role of social image incentives and intrinsic attitudes towards sexual harassment for men (Bénabou and Tirole, 2006), I examine how the training could reduce both harassment and relationships. The training can increase social image incentives against signalling a "bad" type for men since they were trained with their peers apart from changing their intrinsic attitudes towards sexual harassment. Both mechanisms can lead to a reduction in sexual harassment, but only the former can explain the significant decline in relationships. This happens because men might superficially adopt "good" behaviors to avoid social disapproval, leading women to be more cautious in forming relationships with them in the framework. I show the results are consistent with a stronger

change in men's perception of social disapproval than a change in their own intrinsic attitudes which is supported empirically in the data.

Using mechanisms outcomes data, I find that men's perception of their peers' disapproval of sexual harassment increased by 0.056 sd. In addition, men believe their female classmates are more likely to report harassment to their peers rather than to the college's legal complaints committee. However, I find little evidence of a change in men's intrinsic attitudes toward sexual harassment in both the short and long term.⁷ Further evidence from a lab-in-the-field experiment indicates that training men affects how women in their classroom interact with them, consistent with women's increased caution. Specifically, women in treatment classes were 13 percentage points more likely to cooperate with other women compared to control classes.

Other mechanisms seem unlikely. I find that women's relationships only decline with men in their own class but increase with men from outside the class ruling out a complete shut down of women's interactions with men. Additionally, the training did not affect men's choices in the lab experiment or their attempts to form relationships with women, indicating that men were not strongly reducing their approach to women.

Long-run results show that women continued to report lower overall sexual harassment although the results are noisier and insignificant. One key factor that could explain this fading out effect is the disruption caused by the COVID-19 pandemic. In particular, it resulted in the need to recall information about episodes of sexual harassment that occurred between January and March 2020 during the period between March 2021 and April 2022. This might have led to under-reporting in the control group due to recall bias, making it difficult to identify treatment effects. Consistent with this, treatment effects (0.09 sd) for the pre-COVID period can be seen among women surveyed soon after lockdown compared to those surveyed later. Despite these fading effects on harassment, women consistently reported a 47% reduction in romantic relationships with men in the treatment classes, indicating long-lasting effects on opposite-sex relationships.

Lastly, I found no treatment effects on the test scores or psychological well-being of men or women. This suggests that the treatment did not adversely affect academic performance despite reduced relationships. Additionally, there was a significant 0.06 sd increase in women's labor market engagement driven by higher labour market search activities. In a cost-benefit exercise, I show that the training is 12 to 503 times more beneficial than the cost incurred to implement it under various assumptions. Finally, one

⁷This does not rule out that there were smaller positive effects on attitudes of men, however theoretically, these effects could not have been stronger than those on perceived peers' attitudes.

concern is the potential of spillovers of the training from men in the treatment classes to those in control classes which I rule out using student's baseline networks.

To the best of my knowledge, this paper is the first to provide quantitative evidence on the effect of in-person sexual harassment awareness training, using an innovative class-based design tightly linking it to sexual harassment outcomes of women. By providing evidence on how this training works and studying its impact not only on sexual harassment but also on other types of relationships, the paper contributes to future policy making and design of such interventions in similar settings like workplaces or educational institutions. A nascent literature has begun to study ways to deter sexual harassment, such as street patrolling ([Amaral et al., 2023](#)). Very few studies explore sexual harassment in college campuses, with the exception of ([Lindo et al., 2018](#)) who study effects of college partying on the most extreme forms of sexual violence, namely sexual assaults, on campuses. Moreover, some studies in social psychology examine the impact of sexual harassment training on self-reported attitudes ([Roehling and Huang, 2018](#); [Antecol and Cobb-Clark, 2003](#); [Bingham and Scherer, 2001](#)) but cannot causally track the effects on the incidence of sexual harassment. To my knowledge, this is the first paper that fills this gap too.

Most other studies focus on consequences of intimate partner violence (IPV), with far fewer studying active deterrence tools. Sexual harassment includes wide range of behaviours, from sexual remarks or jokes to sexual assault from a stranger or an intimate partner. This necessitates studying different tools and mechanisms for its deterrence. This paper shows that one such important mechanism is the perception of peers' attitudes in interpersonal environments where repeated interactions between potential victims and perpetrators is common. Some deterrence mechanisms studied for IPV are women justice centers ([Sviatschi and Trako, 2024](#)), female management ([Adams-Prassl et al., 2022](#)), unemployment insurance ([Bhalotra et al., 2021](#)) and job offers for women ([Kotsadam and Villanger, 2020](#)). Others study factors such as cash transfers, gender wage gaps, female labor force participation and its consequences for victims ([Bloch and Rao, 2002](#); [Aizer, 2010](#); [Anderberg and Rainer, 2013](#); [Erten and Keskin, 2018](#); [Calvi and Keskar, 2021](#)).

Moreover, this paper sheds light on deeper mechanisms that are invoked by a group-based awareness training. It shows that awareness interventions can be effective by changing individual's perceptions of the others they are trained with. Some studies have focused on changing the attitudes of women's communities or families such as ([Abramsky et al., 2014](#); [Dean and Jayachandran, 2019](#); [Banerjee et al., 2019](#); [Green et al., 2020](#)), while others have aimed to change direct gender attitudes for women

empowerment (Dhar et al., 2022). While none of these studies target sexual harassment, in their secondary results, (Dhar et al., 2022) show that there is no reduction in men’s self-reported perpetration or women’s reporting of sexual harassment despite improved gender attitudes in a classroom-based gender attitudes intervention. This suggests that perception of peers’ attitudes might be more important for changing behaviour like sexual harassment than own intrinsic attitudes.

This paper, thus, also contributes to the literature on social image concerns, showing that group based awareness interventions can activate these concerns by creating common knowledge about others’ awareness or attitudes, which can help deter undesirable behaviours such as harassment, bullying, racist remarks among others. The literature shows that perceptions of what others think can drive one’s behavior, whether the perceptions are accurate or not (Bénabou and Tirole (2006); DellaVigna et al. (2012); Bursztyn and Jensen (2015); DellaVigna et al. (2016); Bursztyn et al. (2020a); Bursztyn et al. (2020b); Bursztyn and Yang (2022)).⁸ The paper further traces down the effects of these changed perceptions of others, for both sexual harassment and another auxiliary outcome—opposite sex relationships—that helps understand the potential of this training in other settings, such as workplaces.

The rest of the paper is structured as follows. In Section 2, I discuss the context of the study. Section 3 details the intervention, Section 4 discusses the experimental design, Section 5 presents the results on main outcomes and mechanisms. Section 6 presents the theoretical framework, Section 7 discusses alternative mechanisms, Section 8 shows impact of the female intervention and Section 9 discusses welfare implications. Section 10 concludes.

2 Context

The United Nations Development Programme defines sexual harassment as “any unwelcome sexual advance, request for sexual favour, verbal or physical conduct or gesture of a sexual nature, or any other behavior of a sexual nature that might reasonably be expected or be perceived to cause offence or humiliation to another person.” The Indian law identifies it as “any unwanted or unwelcome behaviour of a sexual nature” (Sexual Harassment of Women at Workplace Act 2013).

⁸The importance of perception of peers attitudes or class culture in effectiveness of the training is in line with other scholarly recommendations in social psychology and management in how organizational culture is critical for effectiveness of sexual harassment training although none of these measure sexual harassment (Zelin and Magley, 2021; Cheung et al., 2018; Roehling and Huang, 2018).

I adapted the *Sexual Experiences Questionnaire (SEQ)* widely used in social psychology to measure sexual harassment in colleges and workplaces (Fitzgerald et al., 1995; Fitzgerald, 1988). I asked women at baseline about their exposure to different types of sexual harassment two months prior to the survey. The incidents were categorized into mild, intermediate, and extreme which has been previously used for classification by the US Merit System Protection Board (USMSPB, 1981, 1987) and suited this context too. The prevalence of sexual harassment is high as depicted in Table 1: 44%-47% of the women experienced mild and intermediate incidents, while 16% of women reported extreme events. Additionally, 12% encountered mild events, 8% intermediate, and 3% extreme, often from someone within own college as shown in appendix figure C.1, which can be particularly damaging due to its repeated nature. This paper will focus on incidents originating from within the college.

Table 1: Prevalence of sexual harassment at baseline

Variable	Mean	Std. Dev	Min	Max	N
Low-intensity events	0.44	0.49	0	1	1,201
Intermediate-intensity events	0.47	0.49	0	1	1,202
Extreme-intensity events	0.16	0.36	0	1	1,189

Note: The table reports the baseline prevalence rate of sexual harassment. Female students were asked about their exposure to different types of sexual harassment events in the two months before the survey. Mild events include sexual remarks, jokes, and being asked repeatedly out on a date; intermediate events are physical intimidation, stalking, staring, and online sexual harassment; and extreme events include sexual assault and physical contact without permission such as groping, pinching, and fondling. N varies as per how many women agreed to report that type of harassment.

To understand, baseline student characteristics that are correlated with higher reported sexual harassment, baseline data sexual harassment is regressed on various class characteristics as show in Table C.1. It reveals significant correlations between women’s self-reported harassment and their perception of classmates’ support for harassment victims, as well as their awareness of the college’s harassment reporting mechanisms (internal complaints committee). Both are expectantly negatively correlated with reported sexual harassment. Indeed, perception of informal support is a key mechanism that influences the effectiveness of the training as discussed in subsequent sections.

3 Details of the intervention

I summarize the detailed training content in table C.2. Using Safecity's existing program, we tailored the sexual harassment awareness training for college students. It included two main components: awareness and empathy building. The awareness component informed men about the legal aspects of sexual harassment under the Indian Law, principles to detect sexual harassment, and included exercises based on hypothetical situations. These scenarios were first reviewed by students, then discussed in the context of the law by trainers during the sessions, promoting interactive rather than traditional lecture-based training. The training content, including hypothetical situations, was developed with Safecity and legal experts experienced in sexual harassment. This part of the training clarified the legal definition of sexual harassment to help men identify and understand it. Discussions covered various types of sexual harassment, with the aim that this knowledge would alleviate potential awareness constraints for men.

The second component focused on empathy building and behavioral change, aiming to help men grasp the long-term impact of sexual harassment on victims. Men were presented with anonymous narratives from female victims within their courses to facilitate discussions on the effects of behaviors classified as sexual harassment (see appendix B.6 for an example). This approach aimed to build empathy and reduce backlash against the training. A volunteer male student read the narrative followed by a discussion led by Safecity's trainers. Since the training aimed at building empathy for victims, this also reduced the chances of any backlash against the training itself (Bingham and Scherer, 2001).

The training consisted of two sessions: a 90-minute workshop followed by a men-only follow-up session to clarify any doubts. Led by both male and female trainers, the sessions included role-playing and small skits to encourage discussion. The training concluded with strategies for men to intervene and help prevent sexual harassment.⁹ An exhibit of the training is provided in appendix figure C.3.

The female awareness intervention included only the informative component without any discussion with the NGO trainers, but undertaken in the same in-person group setting as the male awareness intervention.¹⁰ Remember this is the same information

⁹Male training occurred in complete privacy and isolation, giving men a protected space to discuss their views openly. Feedback from trainers post-training revealed that men liked the candid nature of this training.

¹⁰Piloting and discussion with NGOs revealed that women understood the concepts quickly and empathy building was much less relevant for them. They were given the same information to read, think about, answer questions on and could clarify any doubts from trained project team members.

that was provided to all women in male intervention colleges. Thus, female awareness intervention college helps to study the independent effect of this information.

4 Experimental design

4.1 Recruitment and randomization

The surveys were conducted by a team trained by Abdul Latif Jameel Poverty Action Lab (J-PAL, South Asia) at the South Asia Center. We coordinated with faculty to schedule class slots for the survey, which were not announced in advance to avoid influencing student attendance. We surveyed students present on the designated days, mitigating selection bias concerns for training or surveys. Both male and female students were informed about the research collaboration between J-PAL and multiple colleges, given an overview of the project's goals, and asked for their informed consent before participating.

The male intervention, was a class-based randomization, where a class was a combination of course, year, and section.¹¹ Classes were stratified according to year of study, field of study, and sex ratio to provide the male training. The sample was divided into 25 strata, and the distribution of classes for each strata is available in Table C.3. Concurrently, all women received the informative component of sexual harassment as well. This was done to remove any under-reporting due to gaps in women's understanding about what constitutes sexual harassment.

Class was a natural unit of randomization. Such training is usually delivered in groups rather than individually when delivered in person. The project's main goal was to examine whether increasing men's awareness reduces women's exposure to sexual harassment. Implementing the training at the class level allowed me to avoid the difficulty of asking women about men from their own class without revealing their identity. Further, the group format fostered greater interaction and maintained confidentiality, enhancing the effectiveness of the sessions.

The female intervention involved two levels of randomization. First, classes were stratified by year, field, medium of study, and sex ratio. Then, 69 classes were categorized into high- and low-intensity classes, where 75% and 25% of women respectively

¹¹It also included a medium if the classes were divided by one. Course means the core subject (like economics and math), which could belong to different fields like science, the humanities, or commerce. Sections were usually created for courses with high demand, and a course had a maximum of three sections. Medium refers to the language of instruction, which could be Hindi or English. Year was the year of study, which could be the first, second, or third year.

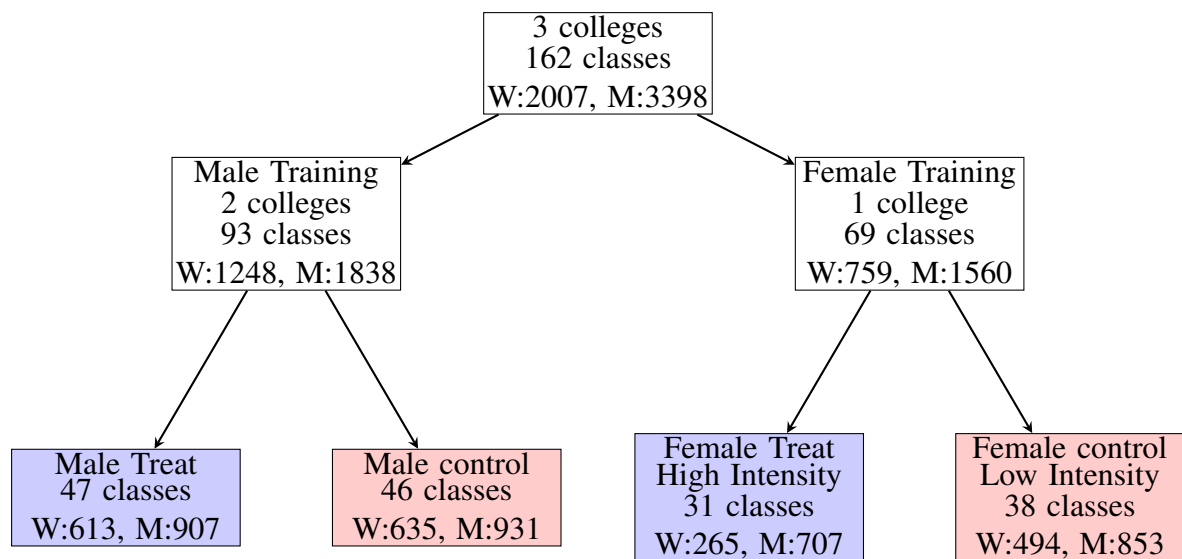


Figure 1: Overall design

The figure above shows the overall design of the experiment. It specifies the number of colleges, classes and male and female students under male and female training.

received sexual harassment information. Women within a class then were randomly selected within the class to receive this information. Figure 1 shows how classes were divided between treatment and control for both the interventions. The experimental variation in intensity across classes for female intervention was done to understand whether the treatment effects on the treated were affected by the proportion of their treated peers. In analysis for the female intervention, I compare high intensity to low intensity classes in a full specification with individual level randomization. The latter helps understand whether women’s reporting changes due to a change in the proportion of her peers treated while keeping her own awareness constant to remove any reporting effects.

After introducing the project, the survey team separated female students into a private room while male students remained in the original classroom. Surveys were conducted digitally on tablets, ensuring privacy and preventing influence from peers. All students consented on the tablets before participating, with less than 3% refusing.¹² In the male intervention, men in treatment classes received information about the project and were told that Safecity wanted to discuss their perspective on sexual harassment. Women in the female intervention received similar information on their tablets.

¹²For female surveys, I also provided a helpline number and Safecity’s helpline.

4.2 Sample, timeline, and balance tests

The two colleges selected for the male intervention comprised 93 classes, with 47 randomly chosen to receive training. The survey timeline is detailed in Figure 2. Baseline surveys and training occurred early in the academic semester, followed by the first endline survey at the start of the next semester. The second endline was conducted over the phone due to lockdowns or after students had graduated.

Since the intervention and the endline survey were sufficiently spaced apart within the academic semester, Hawthorne effects—where behavior changes due to awareness of being observed—were unlikely. Neither students nor faculty knew about the two planned survey rounds. Women were asked to recall men’s behavior from the period between the intervention and endline, making it improbable that men altered their behavior in anticipation of the re-survey.

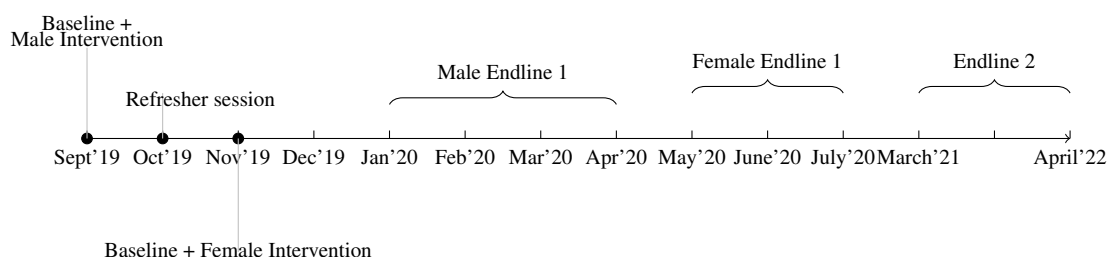


Figure 2: Timeline

A total of 3,086 men and women took part in the surveys, 1,248 women and 1,838 men. Table 2 presents the balance tests indicating balance on all features. Most participants have highly educated parents, 36% belong to historically disadvantaged groups, and about 75% live with their family, and 23% have a working mother. The F-statistic for joint significance is 1.07 with a p-value of 0.38, allowing rejection of the hypothesis that all variables jointly explain assignment to treatment. Appendix Table C.4 presents the balance tests separately for men and women. In both groups, all the socioeconomic demographic variables are balanced except for women’s caste which is statistically expected given the large number of baseline variables.

The survey team reached 83% of the female and 80% of the male baseline populations for the first endline for male intervention, achieving an overall coverage of 82% for the male intervention. They also included additional students not present at baseline, enhancing statistical power. For these extra students, baseline variables recorded at endline showed balance as indicated in appendix table C.5. For long run, the survey team was able to reach 65% of the baseline sample. Tables C.6 and C.7 show there was no differential attrition by treatment status or baseline controls in the short or the long

Table 2: Balance tests for women and men at baseline for male intervention

Control Variable	Treatment Mean	Control Mean	N	p-value
Father education primary	0.05	0.07	2,454	0.16
Father education secondary	0.26	0.28	2,454	0.28
Father education higher	0.68	0.64	2,454	0.16
Mother education primary	0.13	0.14	2,413	0.41
Mother education secondary	0.27	0.31	2,413	0.16
Mother education higher	0.58	0.54	2,413	0.18
Proportion general caste	0.64	0.62	2,675	0.30
Proportion SC/ST/OBC*	0.36	0.37	2,675	0.30
Proportion other groups	0.01	0.01	2,675	0.52
Living in PG/hostel/flat	0.26	0.25	2,675	0.89
Living with family	0.74	0.75	2,675	0.89
Working mother	0.22	0.23	2,902	0.75
Homemaker mother	0.44	0.44	2,902	0.93
From Delhi	0.62	0.61	3,086	0.64
Number of classes	47	46		
Number of students	1,520	1,566		
F-stat	1.07			

Note: The table reports the mean of baseline characteristics for both men and women in the treatment and control classes for male intervention. It also provides p-values from regressing the characteristics on the class-level treatment variable. Strata and college fixed effects are included, and standard errors are clustered at the class level. *SC/ST/OBC represent castes in India. The p-value for the joint test of significance is 0.38. Number of observations vary because students chose 'I prefer not to answer' option for some demographic questions.

run. The balance table for female intervention shows balance for all variables except 3 out of 16 variables in table C.8. Appendix Tables C.9 and C.10 show there is no differential attrition by treatment and no differential attrition by majority of the baseline controls in either the short or the long run.

4.3 Representativeness and external validity

Students in these colleges come from all over the country, reflected in Figure 3 with 40% of the students from outside of Delhi. Using a proxy of quality, Figure C.2 shows that colleges in project sample are right in the middle of the quality distribution and are quite representative.

Appendix Table C.11 provides a comparison of colleges in the sample with those in India and in Delhi using AISHE 2015-16 data, separately. Colleges in the project sample have similar student-level characteristics to student population in the other two samples except for a slightly lower proportion of female students. Colleges in the project sample have a slightly higher pupil-teacher ratio and a higher proportion of assistant professors compared to others reflecting they were newer. Additionally, they have a similar proportion of female teaching staff but a lower proportion of non-teaching female staff. However, there is no particular pattern to highlight selection in either direction. To generalize the results discussed later, I follow [List \(2020\)](#) and [Holz et al. \(2020\)](#) and report the SANS conditions in Appendix B.7.

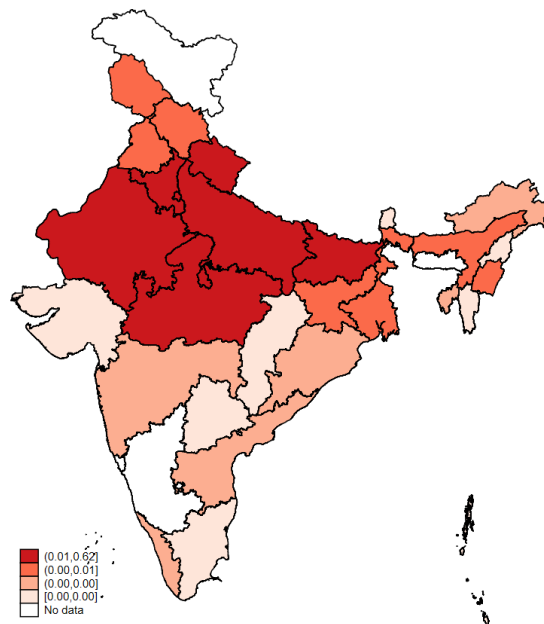


Figure 3: Geographic representation of the students

4.4 Data

4.4.1 Measuring sexual harassment

I had to rely on measuring sexual harassment using survey data since there was no legal reporting done in the college before this project although a legal complaints committee existed in the college. Given the sensitivity of the topic, I undertake a number of precautions in data collection. I used the Sexual Experiences Questionnaire (SEQ), which includes 17 items categorized under gender harassment, unwanted sexual attention, and sexual coercion, grouping some questions by severity. I asked women about sexual ha-

rassment from men in their own class using the SEQ (the final questions are provided in Appendix B.1).¹³

An advantage of asking women about sexual harassment from men in their own class is that it reduces measurement error since women are more likely to recognize their classmates. As mentioned before, I undertook measures to minimize concerns over selection into sample and to ensure women’s awareness was differential between treatment and control classes.

The SEQ minimizes detection differences due to its objective question format, although stigma may still cause under-reporting in treatment classes. [Cullen \(2020\)](#) found no significant differences in reporting of non-partner sexual violence between list and direct methods, similar to those used here. I also run placebo exercises to show that stigma is not driving the results on sexual harassment reported in Section 5.1. Chances of backlash are reduced since women are not asked for perpetrator identities. Lastly, surveys were conducted in a private room with individual tablets, overseen by trained female surveyors to maintain privacy and encourage open responses.¹⁴ While not fool-proof, these practices are more rigorous and comprehensive than those typically found in sexual harassment research and are even more thorough than methods used by the Demographic and Health Surveys for intimate partner violence data collection ([Aguilar et al., 2020](#), [Folke and Rickne, 2020](#), [Kondylis et al., 2019](#)).¹⁵

4.5 Econometric specification

The main econometric specification for understanding the effect of male sexual harassment awareness training on outcomes for both men and women is

$$(1) \quad Y_{icg} = \beta_1 T_{cg} + \beta_2' X_{icg} + \beta_3' K_{cg} + \alpha_g + \gamma_s + \epsilon_{icg},$$

¹³Furthermore, these questions make the answers objective without relying on women’s awareness of sexual harassment. In particular, the questions asked women whether a particular incident “XXX” happened to them rather than asking whether they were sexually harassed, which might be more subjective and prone to information constraints.

¹⁴Female surveyors help in making women participants more comfortable when answering sensitive questions ([Aguilar et al., 2020](#)).

¹⁵Having third-party observers in classes to audit sexual harassment reports was not possible since it would have changed students behaviors. In addition, all women were told that they had the right to withdraw their data if they wanted to, even after submission, and they had the first right over the data that they gave to us. I provided all the women with my contact number and that of a resource person at the University of Warwick in case they wanted to retract their data. This helped to further increase the students’ trust in data privacy.

where i is the student surveyed in the endline survey, c is the class they were in 2019 when baseline was conducted, g is the college student i is in, T_{cg} is whether the class c in college g is assigned to receive the male intervention or not, X_{icg} are student characteristics at baseline, K_{cg} are class characteristics taken from administrative data, α_g are college-level fixed effects, and γ_s are strata (sex ratio \times field of study \times year of study) fixed effects following standard practice (Glennerster and Takavarasha, 2013). Y_{icg} is the outcome of interest. β_1 captures the intent-to-treat effect of the training on student i 's outcome.

The following specification is used for the female intervention results:

$$(2) \quad Y_{ic} = \beta_1 \text{Female_treatment}_{ic} + \beta_2 \text{High_Intensity}_c + \beta_3 \text{High_Intensity}_c \times \text{Female_treatment}_{ic} + \gamma_s + \beta_4 X_{ic} + \epsilon_{ic}.$$

Y_{ic} is the relevant dependent variable, $\text{Female_treatment}_{ic}$ is a dummy that equals 1 if the woman i is assigned to the treatment and equals 0 if she is not, and High_Intensity_c is a dummy that equals 1 if class c was assigned to the high-intensity treatment. β_3 is the difference in the outcome between someone who is treated versus not in the high-intensity class, β_2 is the effect of being an untreated woman in a high-intensity class versus in a low-intensity class, and β_1 gives the effect of being treated in a low-intensity class as compared to someone who is untreated in a low-intensity class. γ_s are strata fixed effects.

In both specifications, standard errors are clustered at the class level, controlling for any correlation in outcomes of students within a class who may be subject to same shocks. Controls are selected by the post-double selection LASSO method (Belloni et al., 2014). If the baseline controls are missing for some individuals, then I control for a dummy variable indicating whether the variable is missing for the respondent or not.

5 Results

5.1 Impact on sexual harassment

As discussed earlier, I use women's reports of different types of sexual harassment perpetrated by men in their own class to study the impact of the sexual harassment awareness training of men. Specifically, I study treatment effects on the proportion of women who reportedly suffered from mild, intermediate and extreme forms of harass-

ment. I also construct a (pre-registered) index called the “same-class index,” which combines all types of harassment together. Moreover, as mentioned previously, these results pertain to all women regardless of whether they were a part of the baseline or not, to improve power (results for baseline women in the appendix).

Table 3 shows the results for the short run in panel A and the long run in panel B. Panel A shows that the male training reduces sexual harassment perpetrated by men from treatment classes by 0.06 sd, as reported by their female classmates about three to four months after the treatment. This seems to be driven by a reduction in the incidence of extreme forms of sexual harassment by 1.05 p.p. (or 0.125 sd) at 1% level of significance. Note that 1% of women in the control group report being harassed physically at least once by men in their class over a period of three months preceding the survey. Thus, the training eliminates the most harmful forms of harassment completely over 3-4 months of the treatment. This also means that the impact on the number of incidents of sexual harassment is likely to be higher. There is also a visible reduction in the milder forms of harassment, however, results are insignificant.

In studying the long-term effects of the treatment, I take into account that the COVID-19 pandemic occurred between the short and the long run endline (which also means that no in-person classes took place after the short-run endline). Hence, women were asked to recall about harassment from before the COVID-19 period. This, as expected, introduces noise into the data. For this, I divided women into those who were surveyed closer to the recall period than the median woman (below median recall) and those surveyed further away from the recall period than the median woman (above median recall). Panel B in table 3 shows the results for these two groups. I find that the women who were surveyed close to the recall period report -0.09 sd lower overall sexual harassment even two to three years after the treatment. Further, the results suggest that there was a decrease in all types of incidents in the long run.

The results are robust to randomization inference, thereby allaying concerns about the low incidence of extreme forms of harassment in the control group for short run. They are also robust to multiple hypothesis testing. In a placebo exercise, I show in Appendix Tables C.12 and C.13 that there are no such negative effects on women’s reporting of sexual harassment from men in a different class or men from outside the college. This shows the results are due to a change in men’s behaviour rather than women’s reporting behavior towards all men. I find a marginally significant increase in women’s reporting of extreme forms of harassment from men outside the college only in the short run. This is because women’s relationships with men outside their college increases after the treatment discussed section 5.3.

The results are also robust to alternative samples over which I created the index, shown in Appendix Table C.14. When I restrict analysis for just the baseline women, in Appendix Table C.15, the results remain qualitatively and quantitatively similar although there is reduced power due to the smaller sample. I find that extreme forms of harassment are lower in the short run, however intermediate forms become higher, and women surveyed closer to the recall period in the long run continue to report lower overall sexual harassment.

Assuming independence of draws in terms of sexual harassment exposure across women across different months, means that for a population of close to 1248 women in my sample, nearly 25 women will suffer from extreme forms of harassment over one semester (6 months). With complete removal of extreme forms of harassment, the estimates translate to 53 fewer women over one academic semester and 159 fewer women over all three years of an undergraduate degree exposed to extreme forms of sexual harassment. This should be taken as a lower bound on the actual number of incidents of sexual harassment since the outcome captures only the proportion of women or the extensive margin.

Overall, there is a strong effect of 0.06 sd to 0.09 sd after a training of three to five hours, comparable with the effects of community-based training programs on IPV. [Abramsky et al. \(2014\)](#) report an effect of -0.31 s.d on physical violence, [Haushofer et al. \(2019\)](#) report -0.2 s.d on sexual violence and -0.16 s.d on physical violence, [Heller et al. \(2017\)](#) report -0.16 s.d on arrests, [Blattman et al. \(2017\)](#) report a null effect on IPV, and [Abramsky et al. \(2014\)](#) report 0.03 s.d on sexual violence of different types of interventions to reduce IPV. With the caveat that these studies focus only on IPV and other crimes, my results for overall sexual harassment are closest to those of [Green et al. \(2020\)](#), who show negative effects of a movie screening program in Uganda on domestic violence reported by women.¹⁶

These results provide the first experimental evidence on the impact of this training for actual sexual harassment. The training for about 3 to 5 hours has a strong effect in the short run and equally big effect in the long run (albeit noisier due to COVID) showcasing the effectiveness of such training. This suggests that academics and policymakers need to be cautious before writing off such training that has been mandated legally in multiple countries ([Dobbin and Kalev, 2020](#)).

¹⁶I first converted these estimates to reflect effects over a three-month period.

Table 3: Women’s self-reported exposure to sexual harassment perpetrated by men in their class

Sexual Harassment	Same-Class Index (1)	Mild Events (2)	Intermediate Events (3)	Extreme Events (4)
<i>A: Short Run</i>				
Male Treatment	-0.062** (0.029)	-0.014 (0.016)	0.011 (0.008)	-0.011*** (0.003)
RI p-value	[0.061]	[0.482]	[0.263]	[0.007]
N	1165	1195	1165	1165
Control mean	0.00	0.07	0.03	0.01
<i>B: Medium Run</i>				
Male treatment	-0.096* (0.053)	-0.028 (0.021)	-0.016 (0.014)	-0.012 (0.009)
Above median recall	0.103 (0.066)	0.012 (0.021)	0.029 (0.021)	0.012 (0.011)
Male treatment × Above median recall	0.024 (0.097)	0.023 (0.036)	-0.017 (0.021)	0.011 (0.016)
F test (Below median = Above median)	0.026	0.122	0.197	0.025
RI p-value	[0.14]	[0.22]	[0.28]	[0.26]
N	699	684	698	699
Control mean	0.00	0.05	0.04	0.01

Note: The main dependent variable is an index of women’s report of sexual harassment from men from their own class (created using Anderson 2008 method). This is regressed on the class level male intervention dummy variable. The index combines questions on different types of sexual harassment perpetrated by men in the same class, as reported by women in columns 2–4. The questions asked female respondents in column 2 if they faced a mild event like sexual remarks, jokes, and being asked repeatedly out on a date; in column 3 if they faced intermediate events like physical intimidation, stalking, staring, and online sexual harassment; and in column 4 if they faced extreme events like sexual assault and physical contact without permission such as groping, pinching, and fondling. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Panel A gives results for endline surveys undertaken between 3 to 4 months after intervention, and Panel B gives gives results for endline surveys undertaken between 2 to 3 years after intervention. The results are medium term since women were asked to recall about sexual harassment from before covid while surveys were done after COVID-19 had started. The above median recall is a dummy=1 iff the woman was surveyed after above median time after recall period had passed. Number of observations vary due to changes in the number of women who choose to report that type of harassment. Randomization inference p-values are reported in square brackets using 1,000 repetitions. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

5.2 Mechanisms

Sexual harassment awareness trainings can affect sexual harassment through various different channels which I discuss next. I will discuss how these outcomes were collected and how they can affect sexual harassment. Later on I will discuss these mechanisms together in a theoretical framework to help interpret these findings.

5.2.1 Are there awareness constraints on men?

I first show that the proof of concept goes through; that is, there are awareness constraints on men about sexual harassment indicating the need for such training. I asked respondents three main types of questions to test awareness: a) hypothetical sexual harassment scenarios (including false positives and detailed in Appendix B.2), b) legal redressal mechanisms, and c) identifying courtship behaviours that are not sexual harassment. These scenarios were developed in consultation with NGO's and legal experts working on sexual harassment. The objectivity of these questions helps to alleviate concerns about demand effects.

Table 4, column 1 presents results on men's awareness combining all awareness questions in an index. I find that men in the treatment classes are 0.09 sd (Panel A, column 1) more aware of what sexual harassment is nearly four months after the training than men in control. These effects persist in the long run at 0.07 sd (Panel B, column 1). This is a strong effect after a training of 3 to 5 hours.

Appendix figure C.4 shows that awareness about legal mechanisms increases by 107% and awareness about ambiguous sexual harassment situation increases by 12% in the short run. This begs the question if increased awareness is the main driver of results. It is, however, inconsistent with the fact that at least in the short run, extreme forms of harassment go down completely. However, men's knowledge about extreme forms of harassment was already high with no treatment effect on its awareness. Although, higher awareness does seem to play some role for milder forms of harassment that go down (insignificantly) in the short run.

5.2.2 Does men's perception of their peers change after the training?

Role of peers or informal institutions has been considered to have a deterrent effect on crime (Nagin et al., 2013). Within colleges, these informal institutions can be particularly strong given the importance of peers in this setting. Peers can impose sanctions or intervene, call out sexual harassment, or support to sexual harassment victims.

I investigate whether men's perception of their peers changes using different types of questions (detailed in Appendix B.5). Men were asked about their beliefs about their peers hypothetical actions if sexual harassment occurs, and also their second order beliefs about their peers' attitudes. In columns 2 and 3 of table 4, I show that men think that victims are more likely to report sexual harassment to their classmates (rather than the legal complaints committee) and that their classmates are more likely to support a sexual harassment victim. The effect on latter is persistent in the long run as well (post graduation).¹⁷

I also asked a sub-sample of men about their perception of what women in their class will do if they suffered from different types of sexual harassment. Treatment persistently makes men think that the women are more likely to report all the incidents to their peers rather than the college's legal committee (appendix tables C.16). Using their second order beliefs, I find that they perceive that their classmates are more likely to term these same situations as sexual harassment (appendix table C.17). After 2 to 3 years, they continue to think that their peers sanction a perpetrator for all types of harassment (column 1,2,3) and will support a victim (column 7, 8) in table C.18. In a sharp contrast, I find no corresponding effects on women's perception of the peers in the long run.

I also collected these men's perception of the legal costs (probability of being reported to their college's complaint authorities and their trust in the committees effectiveness). Given that the training also provided information on laws against sexual harassment and the formal complaints mechanism in their college, the training has surprisingly null effect on perceived legal costs to harassment in column 4 of table 4. This could be if students felt that their formal complaints mechanism was ineffective (proxied by the fact that even after the treatment, the committee hardly received any complaints from students).

Overall, the importance of perception of peers' attitudes or class culture in effectiveness of the training is in line with other scholarly recommendations in social psychology and management in how organizational culture is critical for effectiveness of sexual harassment training although none test for the effects on sexual harassment itself (Zelin and Magley, 2021; Cheung et al., 2018; Roehling and Huang, 2018)

¹⁷For the long run, these questions were asked retrospectively. For example, "If a sexual harassment victim sought support, what percentage of your classmates do you think would have supported them?".

Table 4: Short and Long run effects of the male intervention on mechanism outcomes for men

Mechanisms	Awareness index (1)	Victim reports informally (2)	Peers will support victim (3)	Perceived legal costs (4)	Victim Blaming attitudes (5)	Social Support Exercise (6)
<i>A: Short Run</i>						
Male treatment	0.093*** (0.025)	0.020** (0.010)	0.029* (0.016)	0.018 (0.033)	0.001 (0.086)	-0.002 (0.014)
N	1,904	1,880	1,904	1,887	1,851	93
Control mean	N(0,1)	0.56	0.73	N(0,1)	2.14	0.02
RI p-value	[0.000]	[0.07]	[0.08]	[0.58]	-	[0.89]
<i>A: Long Run</i>						
Male treatment	0.069*** (0.025)	0.009 (0.010)	0.037** (0.016)	0.026 (0.041)	-0.044 (0.084)	0.025 (0.057)
N	1,371	1,349	1,355	1,347	1,340	1,333
RI p-value	[0.01]	[0.370]	[0.03]	[0.55]	-	-
Control Mean	N(0,1)	0.38	0.61	N(0,1)	2.43	0.56

Note: The table reports results from a regression of the dependent variable on the class-level male intervention dummy variable using short and long-run surveys. The dependent variable in column 1 is an index created by combining questions used to test awareness of sexual harassment. In column 2 the dependent variable asks 'What proportion of your female classmates will report to other classmates in case of an SH incident' and in column 3 asks 'What proportion of your classmates will support a sexual harassment victim', in column 4 it is an index created using questions on the perception of legal support for victims. In column 5 it is the coefficient on the interaction between the treatment class and list treatment; this coefficient provides the differential effect of the male treatment on the number of statements that men agree with in short (panel A) and long run (panel B). In column 6, the dependent variable measures the proportion of students who signed up for volunteering for anti-sexual harassment organizations for short run, and measures whether men in treatment class differentially are more likely to donate to anti-sexual assault organization for long run. This variable in Panel A is created by using Google forms floated at the class level, giving 93 classes as the sample size. PDS LASSO method is used to select controls, and strata and college fixed effects are included in the regressions. Randomization inference p values are reported in square brackets except for list experiment outcomes since they are interaction terms. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

5.2.3 Do men's intrinsic attitudes change?

While awareness may not be a binding constraint (as discussed earlier), it is possible that a change in men's intrinsic attitudes against sexual harassment (whether they think sexual harassment is harmful or not or is a victim's fault) might be key in changing men's behaviour. To address experimenter demand effects in measuring attitudes, I used multiple methods. I assume that men's intrinsic attitudes did not change powerfully if they do not empathize with the issue of sexual harassment or blame the victims for sexual harassment after the training.

First, was a list experiment, where a victim blaming attitude statement was masked within a set of four statements and shown randomly to groups of men within the different classes. The statements for the list are in Appendix B.3. The list experiments help to provide plausible deniability since they ask the participant only about the number of statements that they agree with.¹⁸ I study whether the treatment makes men in list treatment support lesser number of statements (which indicates an improvement in attitudes) differentially in treatment classes.

I ran an obfuscation form exercise where the legal complaints committee of the colleges floated a Google form during the short run, inviting students to intern/volunteer for anti-sexual harassment NGOs. These were real opportunities, and the sign-ups were shared with the NGO's as well. I study if the treatment affects the share of students who sign up to volunteer for the NGOs. For long run, I incorporated a donation experiment, where I allowed students to donate a part of their survey incentive to an NGO that fights against sexual harassment. The result on this are shown in panel B, column 6.

Table 4 shows the results for both of these variables in columns 5 and 6. There is no effect on either men's victim-blaming attitudes or the Google form sign-ups in both the short- and long-run. The coefficients are also very small relative to the mean. This is in contrast to [Dhar et al. \(2022\)](#), who find that a gender sensitization program improved students' gender attitudes in India by 0.179 sd. This could be because sexual harassment attitudes may be particularly difficult to change even after awareness changes and because I focus on older students, for whom such attitudes may be less malleable. Moreover, the length of the sexual harassment awareness intervention is the standard length (source: [Safecity](#)). Hence, this denotes that either such training needs to be repeated or it will be effective only through other channels than a change in attitudes.

In the long run, there is an increase in self reported empathy and attitudes towards supporting victims of harassment (Table C.19)(potentially increasing perception of oth-

¹⁸This has been used in [Bursztyjn et al. \(2020a\)](#) and [Dhar et al. \(2022\)](#) to measure stigmatized attitudes. ([Haaland et al., 2020](#)) recommend using list experiments for measuring such attitudes.

ers' attitudes as well), however these, if true, do not translate into a change in more incentivized measures like donations or victim blaming.

These findings are similar to the insight of [Burszty and Yang \(2022\)](#) and [Burszty et al. \(2020a\)](#) that individual perceptions might not correctly reflect beliefs or actual attitudes of those around but can affect one's own actions nonetheless.

5.3 Impact on opposite sex relationships

It is natural to think that when trained on how to think about sexual harassment along with their peers, sexual harassment awareness training might change men's interactions or relationships with women if it changes their perception of what is 'appropriate' behaviour. The Me too movement, for instance, affected women's economic opportunities in academia by making men more cautious in their interactions with women ([Gertsberg, 2022](#)). If this operates through changes in men's perceived social environment, then understanding impact on these opposite sex relationships is crucial to the mechanisms through which this training may or may not work.

To understand this, men and women were surveyed about their romantic partnerships and friendships with their classmates in both short and long run. I create a variable measuring the proportion of opposite-sex friends that they report from their own class and a dummy variable for whether they report dating someone from their own class (this means after graduation in the case of long run surveys). Table 5 reports regression results of the effects on these relationships in columns 2, 3, 4 (short run) and columns 7, 8 (long run) for both men (upper panel) and women (lower panel).

Panel B, column 2 shows that the training reduces opposite-sex romantic partnerships by 1.2 p.p. in the treatment classes, corresponding to a 64% reduction on average compared to the control mean of 2% as reported by women. The coefficient is reassuringly similar for men, although the effect for them is noisier. However, women also report higher dating with men outside the college more in column 3. The results hold true strongly in the long run (after students have graduated) in columns 7 and 8. Column 4 shows a negative but insignificant effect on friendships. These results on romantic relationships persist even two to three years after the intervention, showing that women's beliefs about men in their class did not change. Combining the short run results, I find that there is a 0.13 sd reduction in an index of opposite sex relationships primarily driven by women's choices in the lab-in-the-field game.

Women continue to report around 47% lower romantic relationships with men in their class and continue to report around 40% increase in relationships outside class

indicating almost perfect substitution in the long run. For men the coefficients continue to be negative, and there is no such evidence of substitution as the women. This means there is a potential for these training to have persistently long run effects on relationships and interactions. The fact that women report more relationships with men outside class indicates that women are not reducing their overall interactions with men but with those in their classes.

5.4 Summary of results for male intervention

Overall, I find that the male awareness intervention leads to a reduction in reported sexual harassment by women in their classes both in the short and the long run by 0.06-0.09 sd. Furthermore, I find that there is a persistent increase in men's awareness about sexual harassment, however, this cannot explain the reduction in sexual harassment. The training increases men's perceptions that their classmates (both men and women) are more likely to sanction behaviours that are perceived to be sexual harassment. This holds true for all forms of sexual harassment even 3 years after the training.

Further, this effect is stronger than any change in men's own attitudes towards harassment in the short or long-run. Finally, romantic relationships between treated men and women in their classes goes down while women actively substitute away from these men with men outside their class indicating it is treated men who are driving this substitution.

To understand these results within a framework, I use a simple signalling framework that builds on existing models of social image. While I do not test the model directly, I show that the results above are consistent with the comparative statics of the framework.

6 Theoretical framework

The framework evaluates how training among peers influences men's behavior, impacting sexual harassment and relationships by changing their own and their perception of their peers' attitudes towards sexual harassment. This approach is based on a common framework that explores social image incentives in economic decision-making ([Bénabou and Tirole, 2006](#); [Bursztn et al., 2020a](#)).

Table 5: Short and Long run effects of the male intervention on opposite sex relationships for men and women

	Short run								Long run			
	Survey measures				Lab-in-the-Field				Survey Measures			
	Opposite-Sex* Relationship index (1)	Dating Same Class (2)	Dating Outside class (3)	Opposite- Sex Friends (4)	Switches (Same Sex) (5)	Stays (Opposite Sex) (6)	Dating Same Class (7)	Dating Outside class (8)	Dating Same Class (7)	Dating Outside class (8)	Dating Same Class (7)	Dating Outside class (8)
<i>Panel A: All Men</i>												
Male treatment	0.042 (0.028)	-0.010 (0.007)	0.012 (0.015)	-0.006 (0.016)	0.022 (0.048)	0.018 (0.048)	-0.009 (0.010)	-0.021 (0.018)				
N	1,895	1,539	1,546	1,810	838	531	1,162	1,162				
RI p-value	[0.24]	[0.15]	[0.41]	[0.71]	[0.55]	[0.01]	[0.93]	[0.62]				
Control mean	0.00	0.02	0.10	0.11	0.40	0.61	0.035	0.125				
<i>Panel A: All Women</i>												
Male treatment	-0.134*** (0.033)	-0.012* (0.006)	0.060*** (0.023)	-0.012 (0.208)	-0.141*** (0.028)	0.000 (0.022)	-0.022** (0.010)	0.075*** (0.031)				
N	1,381	1,144	1,146	1,354	555	525	632	632				
RI p-value	[0.001]	[0.11]	[0.02]	[0.60]	[0.03]	[0.99]	[0.04]	[0.03]				
Control mean	-	0.02	0.18	0.15	0.37	0.68	0.035	0.191				

Note: The table reports results from a regression of dependent variable for men in panel A and women in panel B on the class-level intervention for men. The dependent variable in column 1 is an index using the method of Anderson (2008) created from a combination of dependent variables in columns 2–6 excluding column 3. In column 2 the dependent variable is a dummy variable that asks men and women whether they are dating anyone in their own class. In column 3 it asks them about their romantic relationships outside their class. In column 4 it is the proportion of opposite-sex friends from same class. In column 5, it indicates whether the student switches from the same-sex partner from their own class or not and in column 6 it indicates whether the student stays with the opposite-sex partner from their own class or not. Note that the number of observations for columns 5 and 6 are less because the lab in field is a between-subject design. Values are thus imputed using the KLK method for those who were not in a particular group. Columns 7 and 8 show results for dating for the long run after the students have graduated from college. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

6.1 Social environment

Assume men, M , are senders and women, W , are signal receivers and paired randomly with each other. Men can take two actions: sexually harassing b or non-harassing g and they are either bad (B) or good (G). Women decide whether to accept or reject a man's actions, forming relationships only if they accept. Even if rejected, men can still harass. Women aim to avoid B -type men to prevent future abuse. A proportion p of men are B type. The action space for M is $a_m \in \{b, g\}$ and for W it is $a_w \in \{Accept, Reject\}$. Observers (classmates) can impose social disapproval costs (D) on B -types.¹⁹

Both types of M get 0 utility from undertaking b , G types receive a positive intrinsic utility k from doing g , B types suffer a psychic cost c_i if they do g , where $c_i \sim f(\cdot)$ over $[0, \infty)$. Thus, a B -type man is characterized by (t_i, c_i) , where t_i is the broader type B and c_i is the psychic cost of pretending to be of G type. In contrast, a G -type man only has one broad dimension. Women form the same beliefs as their classmates (social environment), denoted as $P(\cdot)$, formed conditional on the actions of men.

Men's utility is characterized by

$$U(t_i, a_i) = \underbrace{I(W \text{ accepts } a_i)}_{\text{Pairing utility}} - \underbrace{c_i I(a_i = g, t_i = B)}_{\text{Psychic costs for B types}} + \underbrace{k I(a_i = g, t_i = G)}_{\text{Intrinsic utility for G types}} - \underbrace{DP(t_i = B|a_i)}_{\text{Social disapproval}}.$$

The indicator function $I(\cdot)$ equals 1 if an event is true. The first term is the utility from forming a relationship (normalized to 1), the second shows the psychic cost for a B -type man engaging in g , the third term is the intrinsic utility a G -type man gains from g , and the last term represents the social disapproval a man faces if perceived as B -type.

If a woman accepts, she receives u if $t_i = G$, $v(D)$ if $t_i = B$, and 0 if she rejects. I assume that $u > 0 > v(D)$. $v(\cdot)$ is assumed to be a continuous and differentiable function of D and $v'(D) \geq 0$. An increase in disapproval against B -type men decreases a woman's costs from being with a B -type man (e.g., probability of being blamed for sexual harassment if she reports him and costs of reporting a B -type man once she realizes his type could go down). A woman never accepts a man if he undertakes b

¹⁹Even if not observable, I assume that women who are receivers of those actions can tell their peers about actions taken by men towards them. I find empirically that women are more likely to report to their peers about a sexual harassment incident after treatment.

because that is legally sexual harassment, and we assume she is aware of this.²⁰ Thus, the only way a man can match with her is through g . The social environment also holds the same beliefs as the woman.

Timing is as follows i) Nature chooses M 's type with probability p that M is of B type, ii) M takes action a_i toward W : $a_i \in \{b, g\}$, iii) W observes M 's actions and updates her beliefs about M 's type: $Pr(t_i = B|a_i)$ and $Pr(t_i = G|a_i)$, iv) W decides whether she will accept or reject his actions: $a_w \in \{Accept, Reject\}$, v) A relationship is formed if W accepts a_i and not otherwise. The game ends after this. I focus on partial pooling equilibrium with women following a mixed strategy (where they are indifferent between accepting and rejecting) for this game and the definition is provided in the Appendix A.1.

6.2 Equilibrium

W will follow a cut-off strategy. Conditional on any action a_i , she will accept iff $P(t_i = G|a_i) \geq \frac{-v(D)}{u-v(D)}$ and will reject otherwise (accept only if posterior belief is high enough that the man is G). Note that both the social environment and the woman holds the same belief about the level of c^* (where the B -type is indifferent between b and g), conditional on which they update their beliefs. Off-the-path beliefs satisfy the intuitive criterion (Cho and Kreps, 1987).²¹ The following characterizes a partial pooling equilibrium in which a fraction $c^* \in (0, 1)$ of B -type men pool with G -type men and undertake g . The rest separate and undertake b .

Result 1. *There exists a $c^* \in (0, 1)$, st all B -type men with $c \leq c^*$ undertake g and the rest of them do b all G -types undertake g . The social environment and women believe that a fraction c^* of the B -type men pool. In particular, $Pr(t_i = G|a_i = g) = \frac{-v(D)}{u-v(D)}$ and $Pr(t_i = G|a_i = b) = 0$. Thus, the beliefs follow Bayes' rule on the equilibrium path, and a fraction $F(c^*)$ of the B -type men undertake g and the rest undertake b . Sequential rationality then implies that women reject if $a_i = b$ and accept with a probability of q when $a_i = g$.*

²⁰This assumption can be justified because all women in all classes were provided with information on sexual harassment in the baseline.

²¹The beliefs should be such that for any off-equilibrium path information set reached, zero probability should be placed on the types for whom taking the action is equilibrium dominated. Thus, a type will not deviate if the deviation is equilibrium dominated.

Proposition 1. *The male intervention can have two possible effects in the model: increase D (social disapproval costs) or decrease p (percentage of men who are type B). Under certain parametric conditions, the implications of these mechanisms on sexual harassment and opposite-sex relationships are given below:*

1. *More B types pool since it could be more costly to being perceived as a B type: An increase in D increases c^* , thereby increasing the proportion of B -type men who pool with G -type men, which reduces sexual harassment. However, q decreases, leading to a decline in relationships since women accept men's offers less. Overall sexual harassment decreases and relationships decrease.*
2. *More B types transform in to G types: If p decreases, sexual harassment decreases due to a composition effect and also because the remaining B -type men increase pooling (on g). Women's probability of accepting relationship offers when men approach them with g increases since there are more G -type men in their class. This leads to an increase in relationships. As a result, overall sexual harassment decreases and opposite-sex relationships increase.*

I provide the proofs for the predictions above in Appendix A.1. The key intuition is that a shift in D or p can affect B -type's incentive to pool. An increase in D increases this incentive, but women will take that into account, which can reduce q (probability that a woman rejects g) because they suffer from matching with B types. However, q can increase if D increases so much that the woman's dis-utility from matching with B types becomes very small (because $v'(D) > 0$). The overall effect on relationships is, thus, ambiguous for increased in D even though sexual harassment decreases. However, Appendix A.1 shows that under certain parametric conditions, a rise in D leads to a fall in relationships.

Reducing the proportion of B -type men (p) decreases sexual harassment as they now prefer non-harassing behaviors (g), positively affecting relationships since women take these composition effects changes. While both mechanisms reduce sexual harassment, only increased social disapproval (D) may decrease relationships. Increased awareness could further reduce harassment among G -types, as it makes men more conscious of their actions, though B -types are primarily influenced by D .

From the lens of the model, the empirical results for sexual harassment and relationships are consistent with an increase in men's perception of social disapproval (D) more

than a change in men’s intrinsic attitudes. Further, treatment effects on opposite-sex relationships are stronger for female respondents in their first year of college who would have less information about their male classmates, in particular, about men’s ‘types’ (Table C.20) and hence are also more cautious in forming relationships with them. Although, it cannot be fully ruled out that intrinsic attitudes changed (even if not captured by the measures I used), theoretically, the change could not have been stronger than the change in perceived peer attitudes.

7 Alternative Explanations

7.1 Freeze effects on men

I present three pieces of evidence to rule out freeze effects on men’s behavior as the main explanation for the results, where men might become overly cautious in interacting with women due to the training.

First, I conducted a lab-in-the-field experiment where men and women in each class were randomly grouped into mixed or same-sex pairs for a quiz competition. The quiz was a combination of 12 questions from female or male stereotypical domains adapted from the literature on gender stereotypes (Bordalo et al., 2019; Coffman et al., 2019; Coffman, 2014). Participants could choose to stay paired with their partner (stick) or work alone (switch).²² Column 5 in Table 5 shows that the training had no effect on men’s tendency to stick or switch in mixed-sex pairs. Instead, women were more likely to stick with other women in treated versus control. This suggests that men did not become more cautious in interacting with women.

Second, survey data showed no significant change in men approaching women for romantic relationships in both the short and long run (Table C.21). The negative coefficient was less than 10% of the control mean, indicating no significant freeze effect. Third, a mediation analysis shows that romantic relationships cannot mediate the effects on extreme forms of harassment even though the effects on overall sexual harassment decreases and is now insignificant (Table C.22). Moreover, I find that single women too face extreme forms of harassment in control (Table C.23). Thus, dating a man alone cannot predict whether you are harassed by him or not. Additionally, the training had

²²A combination of male and female stereotypical questions makes the quiz gender complementary. Thus the decision to stick or switch cannot be because either thinks that they will be better alone in doing the task. Second, each member of the pair would receive a food coupon or online voucher if they won, ensuring that the decision to stick or switch could not be affected by beliefs that men have a bargaining advantage.

no significant effect on non-extreme forms of harassment, which would have shown stronger effects if men were simply becoming more cautious. Lastly, I show that men reported more direct approaches to date women post-treatment, such as "Directly go and approach her," rather than indirect methods like "Finding out about her through social media and following her in college," which can constitute harassment. This indicates that men did not shift to more subtle, ineffective courtship behaviors.²³

7.2 Reporting effects on women

It is unlikely that women became wary of reporting sexual harassment due to pressure from men after the training. If this were the case, we would expect to see negative effects across all types of harassment, not just extreme forms. Furthermore, the long-term reports of lower harassment after graduation suggest that reporting issues are not driven by sensitivity to reporting. Moreover, as shown before women's reports of extreme harassment from men outside their class did not decrease, indicating that they are not avoiding reporting such incidents. Additionally, men had little reason to dissuade women from reporting, as women were never asked for the perpetrator's identity. Lastly, neither the men nor women knew whether we would be coming to do surveys with them again, and hence such strategic manipulation is difficult especially in the long run.

Concerns that the training might increase perceived judgment for dating within the class do not find support in the data., as women reported no change in perceived judgment. I asked women about 'how likely they would have been judged by their classmates if they formed a relationship with a male classmate', and I do not find that the treatment has any effect on this variable (results omitted for brevity). Further, women increased reporting of romantic relationships with men outside their class which provides evidence against generalized stigma against reporting romantic relationships in this context.

7.3 Spillovers

Finally, there could be concerns about spillovers due to the within college nature of the randomization. There are two main types of spillovers I focus on: spillovers from men in the treatment classes to women in the control classes and spillovers from men in the treatment classes to men in the control classes. I circumvent the former completely

²³Results omitted for brevity.

because I asked women about their own exposure of sexual harassment from men in their own class. This allows me to cleanly measure the effect that the treatment has on women's experiences of sexual harassment from trained men.

For the second type of spillovers, I use information on friendships that I collected at baseline for men in the treatment and control classes. I augment the main specification by controlling for the number of male and female friends that a male student has from outside his own class and for the number of male and female friends he has from the treated classes outside his own. I estimate this for only male students in control classes for male and female friends separately. Appendix Table C.24 shows the results. Reassuringly, I do not find any effects of having male or female friends from other treated classes on men's own awareness, attitudes or perception of their own classmates in the short run. Moreover the coefficients for male or female friends from other treated classes is also small.

8 Comparison with female awareness intervention

I randomized the awareness component to be provided to women in varying proportions in a separate college. So I randomly selected classes to receive high intensity (75% women received the intervention) or low intensity (25% women received it) of the treatment. Interpreting results here needs caution because making women more aware about sexual harassment can create reporting effects but it may also induce changes in women's behavior, affecting their actual exposure to sexual harassment. Disentangling these effects can be challenging and hence the design helps me do so. I compare treated women in high- and low-intensity classes. Hence, I report the F-test of equality of coefficients for these women to understand the effect of the class-level treatment to compare with the male treatment conditional on the women being aware of sexual harassment.

8.0.1 Impact on sexual harassment and opposite-sex relationships

Columns 1 and 2 in Table 6, the full impact for short and long run using both class level and individual randomizations. The short run effect of awareness treatment of women is positive on their reporting of sexual harassment but insignificant (β_1) and the effect fades overtime. Being an unaware woman in a high intensity treatment class shows no detectable effects in short run and long run and most coefficients are close to zero. Finally, the additional effect of being treated in the high intensity class is largely negative in short run and negligible in the long run. To investigate further, I test whether

$\beta_2 + \beta_3 - \beta_1 = 0$, but I do not detect any statistically significant differences between the two. I can rule out an effect of the size of 0.1 or above on sexual harassment with the class-level female intervention.

Table 6: Short and Long run effects of the female information intervention on sexual harassment and opposite sex interactions

	Sexual Harassment		Opposite-Sex	
	Short run (1)	Long run (2)	Short run (3)	Long run (4)
<i>Panel A: All women</i>				
Female treatment β_1	0.07 (0.10)	0.03 (0.13)	-0.09 (0.10)	-0.15** (0.06)
High intensity β_2	0.03 (0.08)	-0.08 (0.08)	0.06 (0.08)	-0.00 (0.08)
High intensity \times Female treatment β_3	-0.18 (0.12)	0.02 (0.16)	-0.15 (0.12)	0.12 (0.11)
N	563	350	595	349
$\beta_2 + \beta_3 - \beta_1 = 0$	0.24	0.71	0.99	0.03
Control mean	0.00	0.00	0.00	0.00

Note: The table shows regression results of estimating equations for the dependent variables on the individual-level treatment for women, the class-level treatment, and the interaction. Clustered standard errors are in parentheses, and strata fixed effects are included in all columns. The PDSLASSO method is used for selecting controls. The results for both the long and short run are shown. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.001$.

Next, column 3 and 4 show the results from a similar specification as above for the opposite-sex relationships. Being a treated woman in low intensity class, leads to a significantly negative impact by 0.15 sd on relationships with men, but this effect is entirely positive for treated women in high intensity classes albeit insignificant. This means that having more treated women around oneself or in ones network almost completely arrests the breakdown in relationships that can occur if you are one of the few who knows.

This gives also more confidence for male intervention results since all women in that case were provided with this same information intervention. So breakdown of relationships is necessarily coming from the male intervention. In the short run, though there were no such differences. Hence, the female intervention does not change women's experiences of sexual harassment but may lead to some break down of relationships only if very few women are treated. Results, thus, show that treating men might be the key to reduce sexual harassment.

9 Secondary Outcomes, Benefit-Costs and External Validity

9.1 Secondary Outcomes

Respondents were asked about various aspects of their well-being, including anxiety (GAD), depression (PHQ), positive and negative affect (PNAS), happiness, self-confidence, and generalized self-efficacy at the endline. This was to determine if the training improved well-being or had any detrimental effects. Students were also questioned about their job or internship experience, search, or aspirations with their college's placement arm. Table 7 shows results for both men and women in the short run for the male training. The treatment does not affect the psychological well-being of either men or women but has a significant impact on women's labor market engagement, which increased by 0.05 s.d. This is due to higher job search activity by women in treated classes. They were more likely to have had an internship experience, searched for a job/internship, or planned to do so. This could indicate empowerment effects on women by training men.

I also studied the impact of the training on students' test scores. Reduced relationships might indicate reduced overall cooperation between men and women, potentially affecting academic performance. I used publicly available test scores data from one of the colleges and combined it with survey data. Students take tests for their 'internal assessment' scores, which account for 25% of their final grade each year. These assessments often require more group assignments than final exams. I collected data from the 2017-2018 to 2021-2022 academic years for both semesters (spring and fall), totaling 10 semester-year combinations.

In the endline estimation sample, 80% of men and 83% of women could be matched to the test score data. I ran the following regression.

$$(3) \quad Y_{icst} = \sum_{t=1}^{10} \beta_t T_c \times I\{T = t\} + \beta'_2 X_i + \delta_c + \alpha_t + \gamma_s + \epsilon_{icst}$$

where i is the student, c is the class they are in, s is elective of the student, and t is the semester-year for which the test score is available. X_i is i 's baseline characteristic, T_c is the treatment status of the class of the student, $I\{T = t\}$ are indicators for semester year t . δ_c , α_t , γ_s are class, semester-year and elective course FE. Y_{icst} is the proportion

of marks out of maximum obtainable marks obtained by the student in that elective. Standard errors are clustered at the class level. Controls are the same as in the main specification. β_t captures the difference in marks obtained by student in treatment versus control for semester-year t relative to the omitted period. The test scores for spring semester of 2018-2019 are treated as the reference category.

The results are shown in figure C.5. The results indicate that there is close to null effect on test scores for men and women (at best show some positive effects for women towards the end)). This allays concerns that the treatment adversely affected academic performance.

9.2 Benefit and Cost Analysis

The full scope of benefits and costs is beyond this paper. I use data from a comprehensive Deloitte report to calculate the costs of workplace sexual harassment and training benefits (Deloitte, 2020). The report provides costs for different types of harassment. By averaging short and long run effects of this paper, I first calculate how many fewer women experience harassment due to the training.

I focus on extreme forms of harassment in the short run for clearer results. These are calculated over different time frames: 3 months, 6 months (1 semester), and 1 academic year, assuming an average rate of decrease for each type. Panel A in table 8 shows the training prevented extreme harassment for about 53 women and milder forms for 77 women over one academic year. Panel B uses these numbers to calculate the training costs per woman avoiding harassment. These estimates likely underestimate the true benefits, as they do not take into account the effects on the intensive margin. The results show that the benefits are 12 to 509 times higher than the costs, even with the most conservative estimate projecting results for only extreme forms of harassment over 3 months (at 126).

9.3 External Validity

I discuss the generalizability of the paper's results through three questions: implications for other settings, frequency of the training and scalability of the training program. First, the training is mandated in workplaces and educational institutions in many countries. If peer interactions and hence their disapproval matters more in some settings, then the training might be even more effective in reducing sexual harassment. On other other hand, it might also reduce opposite gender relationships more if they are indeed formed

in such environments. For instance, importance of workplace to form relationships might be less crucial than universities and hence the negative impact might be low.

Regardless, the results show that over test scores did not get affected (if anything, they might have improved for women) suggesting that such training is less likely to affect the economic outcomes firms or universities care about. Further, the economic incentives in workplaces might be higher to coordinate better which might ameliorate some these effects. The mechanisms also suggest that online training might be less effective as trainees can't gauge others' involvement. Thus, moving to a more digital world might involve rethinking how such trainings are delivered although more work from home or gig work options might reduce in-person interactions for employees as well.

With regards to frequency, in this study, training occurred once per semester with two sessions, which aligns with the typical duration of academic programs. However, in workplaces where employees stay longer, repeated training could lead to a greater reduction in harassment and potentially change men's attitudes and women's perceptions over time reducing the impact on relationships. Lastly, with regards to scalability of the training program, the training was delivered by motivated NGO trainers. Scaling up would require adopting models like "train the trainers" where enthusiastic individuals are trained to deliver the training and address questions. General equilibrium effects would likely reduce overall harassment if training is delivered with peers. The impact on relationships is more complex to predict since reduced interactions in current setting may have reduced opportunities for women to learn about men. If women are able to learn more then the negative effects on relationships might weaken.

10 Discussion and conclusion

This study provides the first experimental evidence demonstrating the effectiveness of sexual harassment awareness training for men in reducing incidents of sexual harassment experienced by women. The training significantly lowers harassment rates both in the short and long term, with large reductions in extreme forms of harassment and also a shift in romantic relationship dynamics between men and women.

It is not surprising that the training does not strongly change intrinsic attitudes towards sexual harassment given its duration. However, given that such training is usually done over a couple of hours rather than over multiple days in general settings (Safecity) plus this is about the only amount of time that universities would voluntarily want to effectively commit to doing it, it means there is a need to either increase the dosage of

Table 7: Effects on secondary outcomes

	Psychological Well Being Index (1)	Labour Market Engagement Index (2)
<i>A: All Men</i>		
Male Treatment	-0.001 (0.012)	0.002 (0.028)
RI p-values	[0.97]	[0.96]
N	1,837	1,864
Control mean	0.00	0.00
<i>B: All Women</i>		
Male treatment	0.010 (0.016)	0.054** (0.024)
RI p-values	[0.14]	[0.22]
N	1,367	1,376
Control mean	0.00	0.00

Note: Coefficients from regression of dependent variable on the class level male intervention dummy variable. Panel A is for Men and Panel B is for women in the short run surveys. Randomization inference p-values are reported in square brackets using 1,000 repetitions. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table 8: Benefit cost ratio analysis

	Milder Events (1)	Intermediate Events (2)	Extreme Events (3)
<i>A: Impact of training over different time horizons</i>			
Averted cases over 3 months	19.2	0	13.2
Averted cases over 1 semester	38.4	0	26.4
Averted cases over 1 academic year	76.8	0	52.8
<i>B: Costs per woman who avoids sexual harassment (USD)</i>			
Projection over 3 months	62.65	-	91.13
Projection over 1 semester	31.32	-	45.66
Projection over 1 academic year	15.66	-	22.78
<i>B: Benefit cost ratio</i>			
Projection over 3 months	11.97	-	125.69
Projection over 1 semester	23.94	-	251.38
Projection over 1 academic year	47.88	-	508.76

Note: Calculations for benefit cost analysis under assumptions discussed in the text.

such training through strict legal enforcement or enhance the way this training is done. One such way as showcased in the paper, is to enhance the perception of social costs to harassing others. This can either be emphasized more in such training to both activate and enhance it for instance using the bystander component of the training.

Moreover, changing attitudes might not help in deterrence. For instance, while [Dhar et al. \(2022\)](#) show that an intense gender attitude training strongly affects intrinsic attitudes of young men and women, it fails to change men's own self-reported sexually harassing behaviours and women's reporting of such behaviour although this was not their focus. [Amaral et al. \(2023\)](#) suggest that attitudes of police officers might be key to understanding what types of incidents they will deter or punish more. In particular, they show that police officers are more likely to condone extreme forms of harassment than others. It is logical to think that men perceive that social disapproval against sexual harassment, in particular extreme forms, might have been activated as a result of the training ([Dobbin and Kalev, 2020](#)).

With the above, the paper adds to the debate on whether these training work. Much of the argument against the training stems from potential of backlash that can come from some 'high proclivity to sexually harassing women types' ([Kearney et al., 2004](#); [Robb and Doverspike, 2001](#)). Training that highlights the role of employees in workplaces as perpetrators can incite severe backlash and stereotypical attitudes, or victim blaming ([Dobbin and Kalev, 2020](#)). The training in this paper focused on inducing empathy and helping men to understand their role as an intervener rather than a perpetrator. For instance, they were asked multiple times about what would they have done in several hypothetical scenarios. This is to my knowledge, the only paper that looks at the impact on sexual harassment of such training, and shows its deterrence effects with a deep dive into its mechanisms.

Indeed the training changed men's perception of their peers permanently and proves to be stronger than any effect on men's intrinsic attitudes to sexual harassment. The training's impact on extreme harassment is strong, with an estimated 52 fewer women affected annually while in the long run, the results on overall sexual harassment persist for women surveyed close to the recall period. However, while the training effectively reduces sexual harassment, women may face challenges in distinguishing between "good" and "bad" men post-intervention. This was consistent with a permanent reduction in romantic relationships but it does not lead to any negative impact on group-based internal assessment scores nor on psychological well-being. On the contrary, women show increased labor market engagement while still in college largely driven by higher search.

Given the relatively brief duration of the training and its low cost, the intervention demonstrates high scalability and potential for significant returns. Indeed, the training proves about 12 to 503 times more beneficial than its costs. Its preventative power is particularly relevant in light of the substantial costs associated with sexual harassment for universities, including legal fees and reduced alumni donations.²⁴ These findings highlight the promise of sexual harassment awareness training in higher education settings and underscore the importance of a group-based approach to inducing behavioral change. Further research is warranted to explore the broader implications of such training, particularly its applicability and effects in workplace environments.

²⁴Campus Sexual Assault Can Cost Universities Millions, Forbes, January 2015.

References

- ABRAMSKY, T., K. DEVRIES, L. KISS, J. NAKUTI, N. KYEGOMBE, E. STARMANN, B. CUNDILL, L. FRANCISCO, D. KAYE, T. MUSUYA, ET AL. (2014): “Findings from the SASA! Study: A cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda,” *BMC medicine*, 12, 122.
- ADAMS-PRASSL, A., K. HUTTUNEN, E. NIX, AND N. ZHANG (2022): “Violence against women at work,” .
- AGUILAR, A. A., E. GUTIERREZ, AND P. S. VILLAGRAN (2020): “Benefits and Unintended Consequences of Gender Segregation in Public Transportation: Evidence from Mexico City’s Subway System,” *Economic Development and Cultural Change*.
- AIZER, A. (2010): “The Gender Wage Gap and Domestic Violence,” *American Economic Review*, 100, 1847–59.
- AMARAL, S., G. BORKER, N. FIALA, A. KUMAR, N. PRAKASH, AND M. M. SVIATSCHI (2023): “Sexual Harassment in Public Spaces and Police Patrols: Experimental Evidence from Urban India,” .
- ANDERBERG, D. AND H. RAINER (2013): “Economic Abuse: A Theory of Intra-household Sabotage,” *Journal of Public Economics*, 97, 282–295.
- ANTECOL, H. AND D. COBB-CLARK (2003): “Does Sexual Harassment Training change Attitudes? A View from the Federal Level,” *Social Science Quarterly*, 84, 826–842.
- AZMAT, G., V. CUÑAT, AND E. HENRY (2020): “Gender Promotion Gaps: Career Aspirations and Workplace Discrimination,” *CEPR Discussion Paper No. DP14311*.
- BANERJEE, A., E. L. FERRARA, AND V. OROZCO (2019): “Entertainment, Education, and Attitudes toward Domestic Violence,” *AEA Papers and Proceedings*, 109, 133–37.
- BATUT, C., C. COLY, S. SCHNEIDER-STRAWCZYNSKI, ET AL. (2021): “It’s a man’s world: culture of abuse,# MeToo and worker flows,” .
- BEAMAN, L. AND J. MAGRUDER (2012): “Who Gets the Job Referral? Evidence from a Social Networks Experiment,” *American Economic Review*, 102, 3574–93.
- BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014): “Inference on Treatment Effects after Selection among High-Dimensional Controls,” *The Review of Economic Studies*, 81, 608–650.
- BÉNABOU, R. AND J. TIROLE (2006): “Incentives and Prosocial Behavior,” *American Economic Review*, 96, 1652–1678.

- BHALOTRA, S., U. KAMBHAMPATI, S. RAWLINGS, AND Z. SIDDIQUE (2021): “Intimate partner violence: The influence of job opportunities for men and women,” *The World Bank Economic Review*, 35, 461–479.
- BINGHAM, S. G. AND L. L. SCHERER (2001): “The Unexpected Effects of a Sexual Harassment Educational Program,” *The Journal of Applied Behavioral Science*, 37, 125–153.
- BLATTMAN, C., J. C. JAMISON, AND M. SHERIDAN (2017): “Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia,” *American Economic Review*, 107, 1165–1206.
- BLOCH, F. AND V. RAO (2002): “Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India,” *American Economic Review*, 92, 1029–1043.
- BORDALO, P., K. COFFMAN, N. GENNAIOLI, AND A. SHLEIFER (2019): “Beliefs about gender,” *American Economic Review*, 109, 739–73.
- BORKER, G. (2017): “Safety First: Perceived Risk of Street Harassment and Educational Choices of Women,” *Job Market Paper, Department of Economics, Brown University*.
- BOUDREAU, L. E., S. CHASSANG, A. GONZALEZ-TORRES, R. HEATH, AND N. B. OF ECONOMIC RESEARCH (2023): “Monitoring harassment in organizations,” Tech. rep., National Bureau of Economic Research.
- BURSZTYN, L., A. L. GONZÁLEZ, AND D. YANAGIZAWA-DROTT (2020a): “Misperceived social norms: Women working outside the home in Saudi Arabia,” *American Economic Review*, 110, 2997–3029.
- BURSZTYN, L., I. K. HAALAND, A. RAO, AND C. P. ROTH (2020b): “Disguising Prejudice: Popular Rationales as Excuses for Intolerant Expression,” .
- BURSZTYN, L. AND R. JENSEN (2015): “How does peer pressure affect educational investments?” *The Quarterly Journal of Economics*, 130, 1329–1367.
- BURSZTYN, L. AND D. Y. YANG (2022): “Misperceptions about Others,” *Annual Review of Economics*, 14, 425–452.
- CALVI, R. AND A. KESKAR (2021): “‘Til Dowry Do Us Part: Bargaining and Violence in Indian Families,” .
- CHEUNG, H. K., C. B. GOLDBERG, E. B. KING, AND V. J. MAGLEY (2018): “Are they true to the cause? Beliefs about organizational and unit commitment to sexual harassment awareness training,” *Group & Organization Management*, 43, 531–560.
- CHO, I.-K. AND D. M. KREPS (1987): “Signaling Games and Stable Equilibria,” *The Quarterly Journal of Economics*, 102, 179–221.
- COFFMAN, K. B. (2014): “Evidence on self-stereotyping and the contribution of ideas,” *The Quarterly Journal of Economics*, 129, 1625–1660.

- COFFMAN, K. B., M. COLLIS, AND L. KULKARNI (2019): “Stereotypes and Belief Updating,” *HBS Working Paper No. 19-068*.
- CULLEN, C. A. (2020): “Method Matters: Underreporting of Intimate Partner Violence in Nigeria and Rwanda,” *World Bank Policy Research Working Paper*.
- DEAN, J. T. AND S. JAYACHANDRAN (2019): “Changing Family Attitudes to Promote Female Employment,” in *AEA Papers and Proceedings*, vol. 109, 138–42.
- DELLAVIGNA, S., J. A. LIST, AND U. MALMENDIER (2012): “Testing for Altruism and Social Pressure in Charitable Giving,” *The quarterly journal of economics*, 127, 1–56.
- DELLAVIGNA, S., J. A. LIST, U. MALMENDIER, AND G. RAO (2016): “Voting to Tell Others,” *The Review of Economic Studies*, 84, 143–181.
- DELOITTE (2020): “economic-costs-sexual-harassment-workplace,” .
- DHAR, D., T. JAIN, AND S. JAYACHANDRAN (2022): “Reshaping adolescents’ gender attitudes: Evidence from a school-based experiment in India,” *American Economic Review*, 112, 899–927.
- DOBBIN, F. AND A. KALEV (2020): “Why sexual harassment programs backfire,” *Harvard Business Review*, 98, 45–52.
- ERTEN, B. AND P. KESKIN (2018): “For Better or for Worse? Education and the Prevalence of Domestic Violence in Turkey,” *American Economic Journal: Applied Economics*, 10, 64–105.
- FITZGERALD, L. (1988): “The Dimensions and Extent of Sexual Harassment in Higher Education and the Workplace,” *J Vocat Behav*, 32, 152–175.
- FITZGERALD, L. F. AND L. M. CORTINA (2018): “Sexual harassment in work organizations: A view from the 21st century.” .
- FITZGERALD, L. F., M. J. GELFAND, AND F. DRASGOW (1995): “Measuring Sexual Harassment: Theoretical and Psychometric Advances,” *Basic and applied social psychology*, 17, 425–445.
- FOLKE, O. AND J. K. RICKNE (2020): “Sexual Harassment and Gender Inequality in the Labor Market,” *CEPR Discussion Paper No. DP14737*.
- GERTSBERG, M. (2022): “The unintended consequences of# MeToo: Evidence from research collaborations,” *Available at SSRN*, 4105976.
- GLENNERSTER, R. AND K. TAKAVARASHA (2013): *Running randomized evaluations: A practical guide*, Princeton University Press.
- GREEN, D. P., A. M. WILKE, AND J. COOPER (2020): “Countering Violence against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda,” *Comparative Political Studies*, 0010414020912275.

- GUTEK, B. A., A. G. COHEN, AND A. M. KONRAD (1990): “Predicting social-sexual behavior at work: A contact hypothesis,” *Academy of Management Journal*, 33, 560–577.
- HAALAND, I., C. ROTH, AND J. WOHLFART (2020): “Designing Information Provision Experiments,” *CEBI Working Paper Series, Working Paper 20/20*.
- HAUSHOFER, J., C. RINGDAL, J. P. SHAPIRO, AND X. Y. WANG (2019): “Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya,” *National Bureau of Economic Research Working Paper No. 25627*.
- HELLER, S. B., A. K. SHAH, J. GURYAN, J. LUDWIG, S. MULLAINATHAN, AND H. A. POLLACK (2017): “Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago,” *The Quarterly Journal of Economics*, 132, 1–54.
- HOLZ, J. E., J. A. LIST, A. ZENTNER, M. CARDOZA, AND J. ZENTNER (2020): “The \$100 Million Nudge: Increasing Tax Compliance of Businesses and the Self-Employed using a Natural Field Experiment,” Tech. rep., National Bureau of Economic Research.
- HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): “The allocation of talent and us economic growth,” *Econometrica*, 87, 1439–1474.
- ILO (2022): “Experiences of violence and harassment at work: A global first survey,” .
- KEARNEY, L. K., A. B. ROCHLEN, AND E. B. KING (2004): “Male Gender Role Conflict, Sexual Harassment Tolerance, and the Efficacy of a Psychoeducative Training Program.” *Psychology of Men & Masculinity*, 5, 72.
- KONDYLIS, F., A. LEGOVINI, ARIANNA AND ZWAGER, AND K. VYBORNY (2019): “Sexual Harassment on Transit: Safe Spaces and Stigma,” *World Bank Working Paper, 9269*.
- KOTSADAM, A. AND E. VILLANGER (2020): “Jobs and Intimate Partner Violence—Evidence from a Field Experiment in Ethiopia,” *Journal of Human Resources*, 0721–11780R2.
- LINDO, J. M., P. SIMINSKI, AND I. D. SWENSEN (2018): “College Party Culture and Sexual Assault,” *American Economic Journal: Applied Economics*, 10, 236–65.
- LIST, J. A. (2020): “Non est disputandum de generalizability? a glimpse into the external validity trial,” *NBER Working Paper*.
- MINISTRY OF EDUCATION, DEPARTMENT OF HIGHER EDUCATION, G. O. I. (2020–21): “AISHE Final Report, 2020-21,” .
- NAGIN, D. S. ET AL. (2013): “Deterrence: A Review of the Evidence by a Criminologist for Economists,” *Annual Review of Economics*, 5, 83–105.

- ROBB, L. A. AND D. DOVERSPIKE (2001): “Self-reported proclivity to harass as a moderator of the effectiveness of sexual harassment-prevention training,” *Psychological Reports*, 88, 85–88.
- ROEHLING, M. V. AND J. HUANG (2018): “Sexual Harassment Training Effectiveness: An Interdisciplinary Review and Call for Research,” *Journal of Organizational Behavior*, 39, 134–150.
- SACERDOTE, B. (2001): “Peer effects with random assignment: Results for Dartmouth roommates,” *The Quarterly Journal of Economics*, 116, 681–704.
- SVIATSCHI, M. M. AND I. TRAKO (2024): “Gender violence, enforcement, and human capital: Evidence from women’s justice centers in Peru,” *Journal of Development Economics*, 168, 103262.
- TALBOYS, S. L., M. KAUR, J. VANDERSLICE, L. H. GREN, H. BHATTACHARYA, AND S. C. ALDER (2017): “What is Eve Teasing? A Mixed Methods Study of Sexual Harassment of Young Women in the Rural Indian Context,” *Sage open*, 7.
- UN (2020): “STEPPING UP TO THE CHALLENGE: Towards international standards on training to end sexual harassment,” .
- ZELIN, A. I. AND V. J. MAGLEY (2021): “Sexual harassment training: Why it (currently) doesn’t work and what can be done,” in *Handbook of interpersonal violence and abuse across the lifespan: A project of the national partnership to end interpersonal violence across the lifespan (NPEIV)*, Springer, 3941–3961.
- ZHU, M. (2019): “Job Networks through College Classmates: Effects of Referrals for Men and Women,” *Working Paper, Syracuse University*, 1.

A Appendix A

A.1 Comparative statics for theoretical framework

Definition 1. A partial pooling equilibrium of this game is characterized by the following: i) The equilibrium strategy of the two types of men: G type ($a_G \in \{b, g\}$) and B type ($a_B \in c$, where $c \in [0, \infty)$); ii) Beliefs of the social environment and the woman, which are given by $P(c^*)$, where $P(c^*) = Pr(t_i = B|a_i)$ and $P(c^*) : \{b, g\} \rightarrow [0, 1]$; iii) The woman's strategy for each action of the man, $a_w : \{b, g\} \rightarrow \{Accept, Reject\}$.

Here, I highlight the proofs for proposition 1. Recall that $c^* = F^{-1}(\frac{-(1-p)u}{pv(D)})$ and $q = \frac{-(1-p)u}{pv(D)} - \frac{v(D)D}{u-v(D)}$.

1. Comparative statics on D:

- $\frac{dc^*}{dD} = F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{v'(D)}{(v(D))^2})] \geq 0$.
- $\frac{dq}{dD} = \frac{dc^*}{dD} + \frac{uv(D)+uDv'(D)-(v(D))^2}{(u-v(D))^2}$.
- Thus if

$$F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{v'(D)}{(v(D))^2})] + \frac{uv(D)+uDv'(D)-(v(D))^2}{(u-v(D))^2} \leq 0,$$

$$\text{then } \frac{dq}{dD} \leq 0$$

Overall, sexual harassment is given by $p(1 - F(c^*))$.

Therefore, $\frac{dS}{dc^*} = -pF'(c^*)\frac{dc^*}{dD}$, and since I established $\frac{dc^*}{dD} \geq 0$, $\frac{dS}{dc^*} \leq 0$.

Relationships are given by $R = q[pF(c^*) + (1-p)]$.

$$\frac{dR}{dD} = \frac{dq}{dD}[pF(c^*) + (1-p)] + qpF'(c^*)\frac{dc^*}{dD}.$$

Relationships fall if

$$F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{1}{(v'(D))^2})] \leq -[\frac{u(v(D)+uDv'(D)-(v(D))^2)}{(u-v(D))^2}][\frac{pF(c^*)+1-p}{pF(c^*)+(1-p)+qpF'(c^*)}].$$

2. Comparative statics on p:

- $\frac{dc^*}{dp} = F^{-1'}[\frac{-(1-p)u}{pv(D)}](\frac{u}{(v(D))^2}) \leq 0$ because $v(D) < 0$.
- $\frac{dq}{dp} = \frac{dc^*}{dp}$, and hence $\frac{dq}{dp} < 0$.

$$\frac{dS}{dp} = 1 - F(c^*) - pf(c^*)\frac{dc^*}{dp} > 0.$$

$$\frac{dR}{dp} = \frac{dq}{dp}[pF(c^*) + (1-p)] + q[F(c^*) + pf(c^*)\frac{dc^*}{dp} - 1].$$

This can be rewritten as $\frac{dR}{dp} = \frac{dq}{dp}[pF(c^*) + (1-p)] - q\frac{dS}{dp} < 0$. This proves the result.

Online Appendices

B Appendix B

B.1 Sexual harassment outcomes collected from the SEQ

The following questions were asked "Did anyone...:

- 1. Comment flirtatiously, make direct or indirect remarks/jokes of sexually suggestive or sexist in nature that made you feel uncomfortable?*
- 2. Try to make unwanted attempts to establish a dating (repeatedly asking you out despite you showing no interest or saying no), romantic, or sexual relationship with you despite you trying to discourage it?*
- 3. Try to get too close to you/try to invade your physical space/brush against you/corner you physically in an intimidating and uncomfortable manner?*
- 4. Try to watch you, follow you from a distance, or stare at you repeatedly, making you uncomfortable?*
- 5. Try to use or display sexual/inappropriate/suggestive material or post vulgar/pornographic/ offensive pictures on messages/email/WhatsApp or made some sexual remark or rumours about you?*
- 6. Try to or attempt to create unwelcome physical contact like pinching, touching, groping, or fondling you (Touching you in areas like thighs, arms, private parts, waist, back, breasts, or your hips) without your permission or consent?*
- 7. Was anyone able to or attempt to fondle, kiss, or rub against private areas of your body, tried to remove your clothes, or put/insert something into your private body parts without your consent?*

Items in 1) to 5) represent gender harassment components of the SEQ, and items in 2), 3), 4), 6), and 7) are part of unwanted sexual harassment. If a student reports that any of this was done by an administrative authority in college, then I refer to it as sexual coercion. For the purposes of this paper, I group 1) and 2) as mild; 3), 4), and 5) as intermediate; and 6) and 7) as extreme forms of harassment. For each of the questions above, I also asked students about the broad category of the perpetrator (someone in their class, someone in same college but not in same class, stranger, administrative member of college, someone near home, other, or I prefer not to answer this question).

For the recall period, the period length differed according to when the intervention was done for one college. For most colleges, I asked for the preceding two months (colleges B, C, and D) at baseline. For endline outcomes, I asked for the preceding three months.

B.2 Hypothetical sexual harassment scenarios for measuring awareness

Men and women were asked whether they think that the three situations below were sexual harassment. They could answer yes, no, and I do not prefer to answer.

- 1. Harish asked Yashika out on a date. She said yes and went out on a date with him. He asked her to go out with him again, but she said no without giving him a reason. Harish got upset about it and asked her why she refused. Yashika told him that she did not think it was fun. Harish agreed and did not ask Yashika out again. Do you think Harish sexually harassed Yashika by calling her to inquire again?*
- 2. Naina and Rahul went out for drinks on a date. Rahul asked Naina if she would like to go dance with him. Naina did join him for the dance. He started touching her physically during the dance, but she thought it was because there was not enough space in the dance floor. So she started to dance a bit further from him, but he would still end up coming close to her. Do you think Rahul was sexually harassing Naina?*
- 3. Ramit, Arun, and Ankur were sitting in the class, making some sexual jokes among each other. The jokes were not pointed at anyone though. Rita and Smriti were sitting in the same room having their lunch but could clearly hear what the guys were talking about. Both of them, however, felt embarrassed and uncomfortable with their conversation but did not say anything. Were Ramit, Arun, and Ankur sexually harassing Rita and Smriti?*

There were two more questions I asked students to test their awareness about the legal complaints committee and their awareness about sexual harassment during courtship.

- 1. Suppose a young man likes another young woman. They do not have many common friends. Which of the following behaviors according to you are acceptable ways for him to approach her or get to know her? (Please select as many as you find acceptable): i) Get her WhatsApp number from common groups and write to her at least 2 or 3 times until she replies, ii) Can stand outside her classes (alone or with friends) to hint that he likes her through indirect comments, iii) Can send one of his friends to go and talk to her, iv) Find out her profile on social media (FB, Instagram, Twitter, etc.) and drop her messages there until she replies, v) Find out where she hangs out (clubs or college societies) to understand her schedule on a usual college day, vi) Find*

out about her by talking to her friends or classmates so that they can tell her that he is interested,
vii) Directly go and talk to her.

2. Which of the following is the internal complaints committee of your college? Options: disciplinary committee, department teacher in charge, the women development cell, internal complaints committee, student union, principal, administrative office, gender sensitization committee, other members of faculty, other.

For the courtship question, I coded the answer as correct if the student did not select options a, b, d, and e. For 2), the correct answer is the internal complaints committee.

B.3 Statements for the list experiment

Students were asked the following: Out of the 3 statements below, can you please tell us how many you agree with? You do not have to state which ones you agree with but rather how many of them do you agree with? Your answer can only be 1, 2, or 3 (**and 4 for the list treatment group**).

1. Sexual harassment of women by men is a manifestation of the fact that men and women are taught to stay away from each other in our society.
2. Restrictions on the mobility of women by parents is valid in the face of increased sexual violence against them.
3. Both men sexually harassing women and women sexually harassing men or sexual violence, in that order, should be punishable by law.
4. **Women who face sexual harassment are usually calling for it.**

List control students received only the first three statements, and list treatment students received all 4. The difference in the number of statements agreed to by the two groups reflects the victim-blaming attitude at the class level.

B.4 Hypothetical sexual harassment situations for measuring attitudes

Men and women were first shown the following three sexual harassment situations:

1. Situation A: Seema was in a college freshers party when one of her male seniors touched her back discretely. No one else saw (it was dark), but she knew. This will be termed as sexual harassment by Indian law.

2. *Situation B: Arun asked Neha out directly for a date first, and she agreed. But she did not find it enjoyable, while he liked it. The second time, however, when he asked her again, she said no. Arun did not know why she said no. So he called her again, but she did not pick up, and he kept calling her multiple times to get an answer. This will be termed as sexual harassment by the Indian law.*
3. *Situation C: Reema was in the bus frequented by other college students as well. She was sitting with one of her male colleagues. He started looking at personal pictures on his phone that were sexual in nature. He kept the phone at a distance, but Reema was able to see what he was watching. This will be termed as sexual harassment by the Indian law.*

The respondents were then asked the following five sets of questions for each of the situations separately:

1. *Please tell us for each of the above situations whether they should be termed as sexual harassment legally.*
2. *Below please tell us for each situation above what percentage of men who are present in your class right now you think will agree that these were sexual harassment.*
3. *Below please tell us for each situation above what percentage of women who are present in your class right now you think will agree that these were sexual harassment.*
4. *Below please tell us for each situation above what percentage of women who are present in your class right now you think will report this incident to the college's ICC if it happened with them?*
5. *Below please tell us for each situation what percentage of women who are present in your class right now you think will report the situations above to other students/teachers/classmates if this happened to them?*

B.5 Measuring perceptions of social and legal costs of sexual harassment

For measuring perceived social costs to sexual harassment, I asked students the following questions: i) What percentage of your classmates do you think will be supportive of you in case you hypothetically wanted to make a sexual harassment complaint against someone?, ii) What percentage of women who are in your class do you think will report or share with other students/teachers/classmates if they were sexually harassed by someone?, iii) Who are the three students you would nominate as someone that students in your class can go to for support or advise for going to the ICC for a sexual harassment incident? I used the proportion of male students from the class, as reported by women, to determine whether it influenced their perception of male support within the class.

To collect data on the perception of the formal costs of sexual harassment, I asked the following: i) What percentage of women who are in your class do you think will report to the college's ICC if they were sexually harassed by someone? and ii) How much do you trust your college's ICC to solve a student's sexual harassment complaint if approached? The options are highly trust them, trust them, trust them a little, and do not trust them at all. I constructed an index for the perceived legal and social costs separately, which I report the results on.

B.6 Anonymized narrative

“This happened when we were all on the dance-floor; everyone was very close to each other, when I felt a hand on my buttock. At first I brushed it off thinking it must have been a mistake; but then it happened again. This time I was sure I was not imagining it. I looked behind, and I am not sure, till date, who this person exactly was. He was definitely a college senior though, because there was a group of 3-4 seniors dancing right behind us. What scares me till date is the fact that I have no idea who this person was. I am still in touch with most of these seniors, and there is a chance that I am still in touch with my own harasser. It happens on a daily basis, sexual harassment, we have normalized it. But a senior from one's own college doing it is something that disgusts me, and frightens me at the same time. I wish I could have said something that day, screamed, anything; but I was so unsure about what I felt myself, I could not have possibly done anything about it.”

B.7 Generalizability

Regarding selection, I collaborated with three colleges, covering all classes, which consisted of 93 classes for the male intervention and 69 classes for the female intervention (discussed in Section 8). For the sample used in the paper, students who were present on the day of the survey (which was unannounced) formed the baseline sample. Students were free to leave the trainings or surveys if they wanted, but less than 5% did so. The surveys took place during regular college hours, which helped me gain access to the student population that would normally be in attendance. When scaling to a bigger population, one must consider that students in these colleges might have been positively selected on household characteristics, ambition, and other factors relevant for admissions in an urban area. I discuss this in Section 4.3.

For attrition, I followed up with 80% of the sample at endline (82% of the women and 77% of men), and there was no differential attrition by treatment status. Most of the students who were not reached (and were supposed to be traced during college hours) could not be easily reached after college closures due to the lockdown.

Regarding naturalness of the choice task and setting, it is important to note that sexual harassment trainings are mandated by law in the educational institutions that I collaborated with on the project. Thus, the setting closely resembles the environment of such educational institutes. Additionally, the set-

ting can be generalized to workplaces due to shared characteristics. Both settings involve repeated interactions between potential perpetrators and victims outside the home, making them suitable for studying the subjects under investigation.

The endline consisted of three types of measures: survey measures, a lab-in-the-field experiment, and a list experiment and Google form data. The survey asked students about their exposure to sexual harassment (due to lack of any naturally occurring data on incidences).²⁵ The lab-in-the-field experiment results were used to understand the patterns in survey measures of inter-personal relationships. List experiments and Google form data were designed in a way that the purpose of the questions or the exercise was not clear to address demand effects. While the list experiment was embedded within the survey, the Google form was floated via the legal complaints committee.

In terms of scaling, certain non-negotiable features include conducting the training for men exclusively and separately, ensuring a consistent timeline for measuring effects, maintaining participants' trust in the safety of their data, and having at least one male trainer present during men's training. Additionally, conducting further replications is important to determine whether providing training to both men and women together yields different effects.

C Attrition and Lee bounds, female information, and difference-in-difference estimates

²⁵It was not feasible to have third-party observers since that would change students' behavior even more.

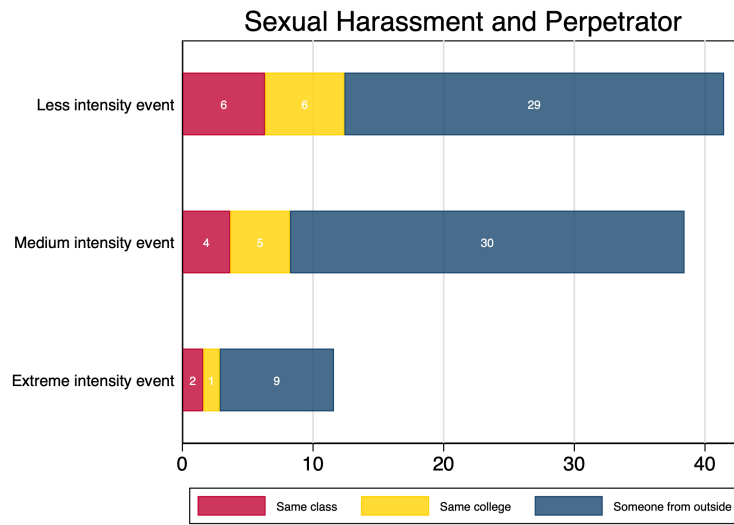


Figure C.1: Perpetrators of sexual harassment as reported by women
 The figure shows the percentage of women who faced sexual harassment from men in different categories. The men are categorized into three groups: a) someone from outside the college, b) someone in college but not in the same class as the female respondent, and c) someone from the same class as the female respondent.

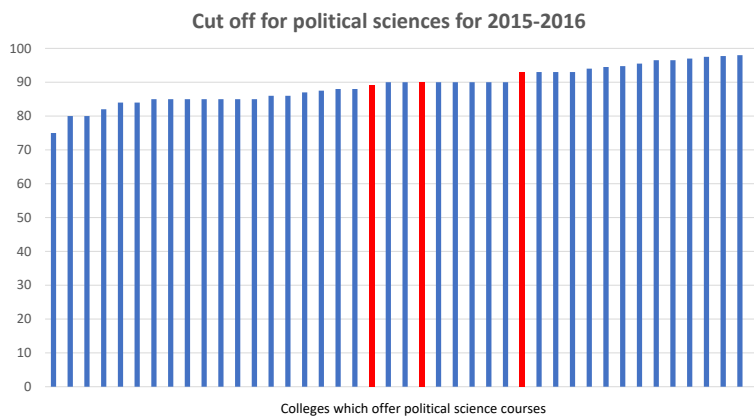


Figure C.2: Quality-wise representativeness of the colleges
 Figure shows the college quality distribution for different colleges in the same university. The colleges studied in this paper are depicted in red. College quality is proxied by admission score cut-offs for entry into the college for the 2015–2016 academic year for an undergraduate degree in political science.



Figure C.3: Male training

Men were taken to rooms with projectors for the training, where two trainers from Safacity delivered the training in a treatment class.

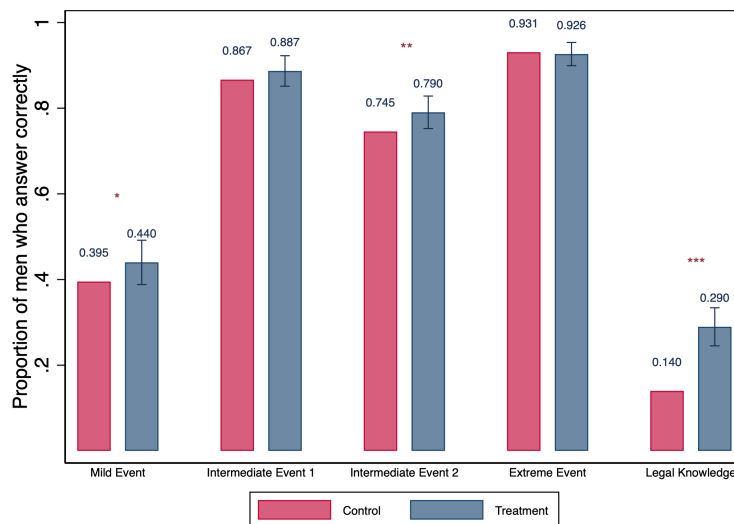


Figure C.4: Treatment effects on men's awareness

The x-axis lists the different types of events on which I tested men's awareness. The y-axis reports the percentage of men who answered the question correctly. Red bars represent the treatment classes, and blue bars represent the control classes.

Table C.1: Correlates of baseline Sexual harassment

Sexual Harassment	(1)	(2)	(3)	(4)	(5)
Perceived informal support	-0.0006* (0.0003)	-0.0007** (0.0003)	-0.0008** (0.0004)	-0.0008** (0.0004)	-0.0007* (0.0004)
Male friends		0.0594 (0.0533)	0.0450 (0.0553)	0.0402 (0.0553)	0.0414 (0.0551)
Mode of transport			0.0065 (0.0092)	0.0085 (0.0091)	0.0087 (0.0091)
Know Complaints committee				-0.0089** (0.0036)	-0.0091** (0.0036)
Self esteem					-0.0399 (0.0266)
N	1240	1159	1091	1091	1091
R^2	0.0091	0.0112	0.0149	0.0206	0.0236

Note: The table reports results from regressing sexual harassment on the class-level variables collected from women at baseline. The independent variables are perceived informal support for victims of harassment, proportion of male friends in the same class, knowledge of the anti complaints committee, Rosenberg self-esteem score, mode of transportation frequently used, whether live in a hostel or not, awareness about sexual harassment (suppressed). The dependent variable takes a value 1 if women report any sexual harassment from men in their class and zero otherwise. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.2: Training content

Basic definitions	<ul style="list-style-type: none"> i) Basic understanding of generalized harassment; ii) Discuss how it is different from sexual harassment.
Reasons for discussing sexual harassment	<ul style="list-style-type: none"> i) Why sexual harassment is being discussed; ii) Discussion of how prevalent men think sexual harassment is; iii) provided with information on the percentage of women in their course of study that reported facing sexual harassment; iv) Discuss the evidence of prevalence from men's own friends, or classmates, seniors or juniors.
Understanding sexual harassment	<ul style="list-style-type: none"> i) What is unwelcomed and unwanted behaviour of a sexual nature; ii) Discussion of a hypothetical situation that students were asked to read on tablets beforehand; iii) Discuss the concept of consent: only yes is a yes, and everything else is a no.
Empathy Building/Perspective taking	<ul style="list-style-type: none"> i) One male student volunteer asked to read an anonymized narrative of sexual harassment incident collected from a female student in their course; ii) Subsequent discussion on the long run impact of apparently innocuous incidents of sexual harassment on victims.
Bystander intervention Do's and Dont's	<ul style="list-style-type: none"> i) Discuss another hypothetical situation that students were asked to read but from an onlookers' perspective; ii) discussion of various ways in which onlookers could intervene; iii) Discussion on reactions if the victim or perpetrator were their own friends.
Part of solution	<ul style="list-style-type: none"> i) Emphasize men's role as part of the solution rather than the problem; ii) Steps or ways for bystander intervention; these included principles of listening and learning from victims, intervening sexual jokes, not blaming victims, not sexually harassing others, intervening after asking victims if they need help.

Table C.3: Number of classes in each strata

Number of Classes	
Strata 1	5
Strata 2	2
Strata 3	2
Strata 4	4
Strata 5	7
Strata 6	3
Strata 7	3
Strata 8	5
Strata 9	2
Strata 10	4
Strata 11	2
Strata 12	3
Strata 13	6
Strata 14	6
Strata 15	5
Strata 16	5
Strata 17	6
Strata 18	5
Strata 19	5
Strata 20	2
Strata 21	4
Strata 22	2
Strata 23	2
Strata 24	2
Strata 25	1

Note: The table shows the number of classes in each strata for the male intervention colleges.

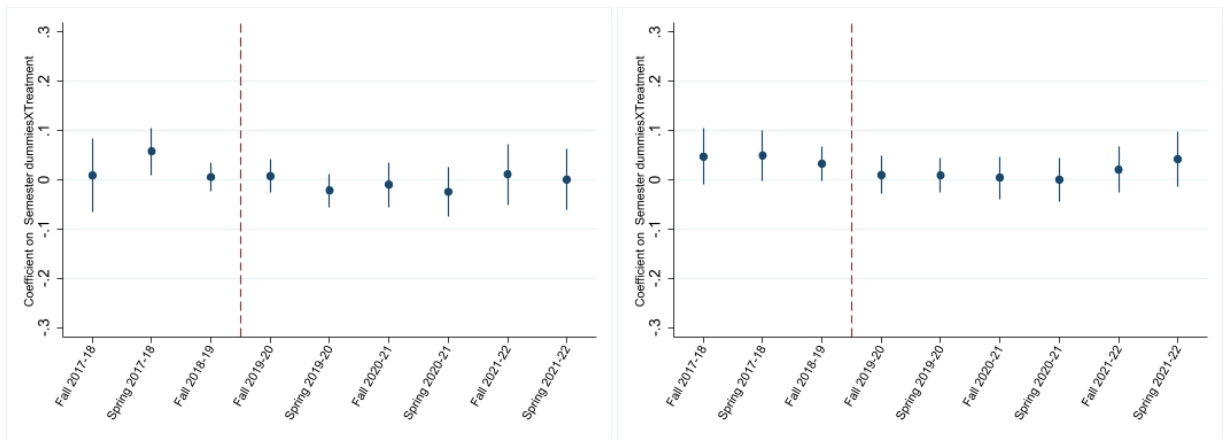


Figure C.5: Coefficient plots for effect of the treatment on test scores for i) men and ii) women

Table C.4: Male intervention: Balance tests for baseline men and women

Control Variable	Men Baseline			Women Baseline		
	β_{Treat}	Control Mean	N	β_{Treat}	Control Mean	N
Father education primary	-0.02	0.32	1,712	-0.03	0.26	1,242
Father education higher	0.04	0.49	1,712	0.05	0.59	1,242
Mother education primary	0.02	0.39	1,702	-0.08	0.34	1,246
Mother education higher	0.02	0.39	1,702	0.09	0.51	1,246
Proportion SC/ST/OBC	0.01	0.39	1,747	-0.06*	0.26	1,248
Proportion general caste	-0.05	0.49	1,747	0.07*	0.64	1,248
Proportion non-Hindu	-0.01	0.10	1,599	-0.02	0.09	1,197
Living in PG/hostel/flat	0.01	0.29	1,517	0.01	0.17	1,244
Living with family	-0.01	0.70	1,517	0.02	0.74	1,244
Working mother	-0.01	0.11	1,747	0.02	0.17	1,248
Homemaker mother	0.00	0.51	1,747	-0.00	0.52	1,248
From Delhi	0.02	0.54	1,827	0.01	0.70	1,248
In Delhi for < 3 yrs	-0.01	0.59	1,827	0.00	0.73	1,248
Mild events				0.01	0.22	1,002
Intermediate events				0.01	0.30	1,007
Extreme events				0.01	0.29	995

Note: The table reports coefficients from separate regressions of baseline characteristic on the treatment status of men both for men and women in the baseline sample. Strata and college fixed effects are included. Standard errors are clustered at the class level.

Table C.5: Male intervention: Balance tests for all endline men and women

Control Variable	Men Endline			Women Endline		
	β_{Treat}	Control Mean	N	β_{Treat}	Control Mean	N
Father education primary	-0.03	0.02	1,749	-0.028	0.035	1,328
Father education higher	0.04	0.04	1,749	0.052	0.042	1,328
Mother education primary	0.00	0.02	1,740	-0.067*	0.040	1,329
Mother education higher	0.02	0.04	1,740	-0.082*	0.047	1,329
Proportion SC/ST/OBC	-0.03	0.03	1,784	0.043	0.031	1,335
Proportion general caste	-0.00	0.02	1,784	-0.018	0.026	1,335
Proportion non-Hindu	-0.01	0.01	1,610	-0.009	0.015	1,259
Working mother	0.00	0.02	1,784	0.02	0.027	1,335
Homemaker mother	-0.01	0.03	1,784	-0.011	0.031	1,335
From Delhi	0.02	0.04	1,850	0.020	0.30	1,335

Note: The table reports coefficients from separate regressions of baseline characteristic on the treatment status of men both for men and women in the final estimation sample collected at endline. Strata and college fixed effects are included. Number of variables in this table are lower than the baseline sample because some variables could have been affected by treatment and hence are not included. Standard errors are clustered at the class level.

Table C.6: Male intervention: Survey attrition at endline and treatment

Control Variable	Covered in Endline1	Covered in Endline2
Male intervention	-0.010 (0.021)	-0.038 (0.029)
Female	0.060*** (0.021)	-0.025 (0.032)
Male intervention \times Female	0.006 (0.032)	0.030 (0.043)
N	3,059	3,059

Note: The dependent variable is a dummy of whether the student was covered in endline or not. The columns report coefficients from the regression on endline coverage on treatment status of men, female dummy, and an interaction of the two. Strata and college fixed effects are included, and standard errors are clustered at the college class level. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.7: Male intervention: Survey attrition at endline, treatment, and baseline controls

Control Variable	Covered in Endline1	p-value	Covered in Endline2	p-value
Father education primary	0.109	0.10	-0.113	0.19
Father education secondary	-0.021	0.58	-0.020	0.68
Father education higher	-0.011	0.74	0.051	0.34
Mother education primary	0.052	0.24	0.060	0.32
Mother education secondary	0.001	0.98	-0.042	0.27
Mother education higher	-0.028	0.40	0.008	0.85
Proportion SC/ST/OBC*	-0.028	0.40	-0.000	0.99
Proportion general caste	0.014	0.66	-0.004	0.91
Proportion other groups	0.095	0.78	-0.031	0.49
Living in PG/hostel/flat	-0.033	0.42	-0.031	0.49
Living with family	0.033	0.42	0.031	0.49
Working mother	0.007	0.81	-0.050	0.22
Homemaker mother	-0.002	0.92	-0.018	0.57
From Delhi	-0.129	0.67	0.032	0.49

Note: The table reports coefficients from the interaction term between the male intervention and the baseline control in a regression of endline coverage on treatment, baseline control, and an interaction of the two. Strata and college fixed effects are included, and standard errors are clustered at the college class level.

Table C.8: Balance tests for women for female intervention (class level)

Control Variable	High treatment intensity Mean	Low treatment intensity Mean	N	p-value
Father education primary	0.015	0.10	2,187	0.19
Father education secondary	-0.007	0.29	2,187	0.69
Father education higher	0.001	0.49	2,187	0.98
Mother education primary	0.018	0.21	2,189	0.20
Mother education secondary	-0.040	0.30	2,189	0.01
Mother education higher	0.025	0.33	2,189	0.22
Proportion SC/ST/OBC*	0.012	0.39	2,262	0.03
Working mother	0.009	0.18	1,584	0.65
From Delhi	-0.046	0.54	2254	0.15
Living in PG/hostel/flat	0.030	0.25	2,639	0.06
Aims to study after college	0.022	0.72	1,976	0.20
Aims to work after college	-0.024	0.24	1,974	0.78
Has undertaken job/internship	-0.004	0.11	2,167	0.42
Will sit for job interviews	-0.021	0.46	1,975	0.95
Sat for job interviews	-0.004	0.06	2,149	0.68
Undertook job through the p-cell of the college	-0.004	0.04	2,088	0.61
Number of classes	31	38		
Number of students	972	1,347		

Note: The table displays regression coefficients obtained from a regression of the variables mentioned above on the treatment status of the student's class. *SC/ST/OBC represent castes in India. Strata fixed effects are included, and standard errors are clustered at the class level. The table includes both men and women in the sample.

Table C.9: Survey attrition at endline and treatment for female intervention

Control Variable	Covered in Endline1	Covered in Endline2
Female treatment	0.02 (0.026)	-0.002 (0.008)
Female	0.176 (0.024)	0.515 (0.029)
Female treatment × Female	-0.043 (0.036)	-0.013 (0.044)
N	2,262	2,262

Note: The dependent variable is a dummy of whether the student was covered in the endline or not. Female is a dummy equal to 1 if the respondent is a woman and 0 if they are a man. Treatment variable is 1 if class assigned to the high intensity treatment. Strata fixed effects are included. Standard errors are clustered at the class level. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.10: Survey attrition at endline, treatment, and baseline controls for female intervention

Control Variable	Covered in Endline1	p-value	Covered in Endline2	p-value
Father education primary	-0.032	0.567	0.053	0.32
Father education secondary	0.038	0.321	-0.007	0.84
Father education higher	-0.047	0.196	-0.015	0.70
Mother education primary	-0.045	0.314	0.065	0.15
Mother education secondary	-0.059	0.148	-0.046	0.27
Mother education higher	0.015	0.718	-0.007	0.85
Proportion SC/ST/OBC*	-0.082	0.053	-0.001	0.95
From Delhi	-0.073	0.045	-0.069	0.035
Living in PG/hostel/flat	0.045	0.374	0.091	0.023
Aims to study after college	0.000	0.998	0.026	0.415
Aims to work after college	-0.025	0.608	-0.034	0.310
Has undertaken job/internship	-0.035	0.538	-0.052	0.24
Will sit for job interviews	0.011	0.791	-0.040	0.30
Sat for job interviews	-0.061	0.475	-0.052	0.24
Undertook job via p-cell	-0.012	0.917	-0.003	0.96

Note: The table reports coefficients from the interaction term between the female intervention and the baseline control in a regression of endline coverage on the treatment (whether in high intensity class or not), the baseline control, and an interaction of the two. Strata fixed effects are included, and standard errors are clustered at the class level. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.11: Comparison of student sample with AISHE 2015–2016 data

Variable	Men and Women		
	Mean AISHE India	Mean AISHE Delhi	Mean Sample Baseline
Panel A: Student-Level Variables			
% Female students	0.48	0.46	0.40
% SC/ST/OBC*	0.53	0.26	0.37
% General caste	0.47	0.74	0.62
% Muslims & other minorities	0.09	0.10	0.10
% Living in hostel	0.16	0.15	0.12
Panel A: College-Level Variables			
Pupil-Teacher Ratio	21	19	25
% Assist. professors	0.66	0.46	0.94
% Associate professors	0.11	0.23	0.05
% Professor	0.09	0.15	0.005
% Female teaching staff	0.39	0.52	0.47
% Female non-teaching staff	0.39	0.29	0.13

Note: Panel A provides a comparison of student-level characteristics between students in all universities and colleges in Delhi, India, and the baseline sample. Panel B provides the same comparison but for college-level characteristics. *SC/ST/OBC represent castes in India.

Table C.12: Sexual harassment reported by women from men from a different class

	Different Class Index			Mild events			Intermediate Events			Extreme events		
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
Male treatment β_1	-0.017 (0.04)	0.032 (0.07)	-0.002 (0.01)	0.006 (0.02)	-0.011 (0.01)	0.015 (0.01)	0.002 (0.004)	0.001 (0.001)				
Above median recall β_2	-	-0.02 (0.06)	-	0.012 (0.04)	-	0.01 (0.02)	-	-0.01 (0.001)				
Male treatment \times Above median recall β_3	-	0.02 (0.12)	-	0.02 (0.05)	-	0.01 (0.02)	-	0.01 (0.01)				
N	1,288	753	1,195	684	1,165	698	1,267	745				
Control mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				

Note: The main dependent variable is an index of womens report of sexual harassment from men from different class (created using Anderson 2008 method). The questions asked female respondents about different types of sexual harassment perpetrated by men from a different class in their college. The events are same as in the main results table on sexual harassment. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Asterisks denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Table C.13: Sexual harassment from men outside the college

	Outside Class Index			Mild events			Intermediate Events			Extreme events		
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
Male treatment β_1	0.048 (0.04)	0.001 (0.06)	0.031 (0.02)	-0.023 (0.04)	0.055 (0.03)	0.003 (0.04)	0.021* (0.01)	0.009 (0.019)				
Above median recall β_2	-	-0.084 (0.05)	-	0.012 (0.04)	-	0.01 (0.04)	-	-0.01 (0.02)				
Male treatment \times Above median recall β_3	-	0.02 (0.09)	-	0.02 (0.06)	-	0.01 (0.06)	-	0.01 (0.03)				
N	1,288	753	1,195	684	1,165	698	1,267	745				
Control mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				

Note: The main dependent variable is an index of women's report of sexual harassment from men from outside the college (created using Anderson 2008 method). The questions asked female respondents about different types of sexual harassment perpetrated by men from outside the college. The events are same as in the main results table on sexual harassment. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Asterisks denote significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Table C.14: Alternative samples for constructing sexual harassment index

Sexual Harassment	SH from Same-Class Index (Only Extreme Events Sample)	SH from Same-Class Index (All-Reported Sample)	SH from Same-Class Index (Simple Addition)
<i>Panel A: No Controls</i>			
Male treatment	-0.062** (0.029)	-0.065** (0.030)	-0.066** (0.031)
<i>Panel B: With Controls</i>			
Male treatment	-0.062** (0.028)	-0.065** (0.030)	-0.066** (0.030)
RI p-values	[0.061]	[0.063]	[0.060]
N	1,165	1,105	1,105
Control mean	0.00	0.00	0.00

Note: The table reports results from regressing the dependent variable (sexual harassment reported by women) on the class-level male intervention dummy variable, estimated only on the sample of women. Column 1 reports results for only the women who reported extreme events, and column 2 reports results for those who reported all events and uses weights from this sample. Column 3 reports results for those who reported all events, but weighting is done using that from the entire sample of women used in the main table. The dependent variable in each column is constructed using the Anderson (2008) method, as in the main table, but the samples differ across different columns. The dependent variable in column 1 is an index created using the Anderson (2008) method for only those women who reported extreme events. In column 2, the dependent variable is only those women who reported all events, and in column 3, it takes the simple addition of components of the index. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Randomization inference p-values are reported in square brackets using 1,000 repetitions. *p<0.1, **p<0.05, and ***p<0.01.

Table C.15: Women's self-reported exposure to sexual harassment perpetrated by men in their class (baseline women)

	Same Class Index		Mild events		Intermediate Events		Extreme events	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
Male treatment β_1	-0.026 (0.031)	-0.087* (0.048)	-0.008 (0.015)	-0.027 (0.021)	0.018** (0.009)	-0.016 (0.014)	-0.006* (0.003)	-0.012 (0.009)
Above median recall β_2	-	0.095* (0.057)	-	0.012 (0.021)	-	0.029 (0.021)	-	0.012 (0.011)
Male treatment \times Above median recall β_3	-	0.034 (0.087)	-	0.023 (0.036)	-	-0.017 (0.021)	-	0.011 (0.016)
N	968	732	905	684	890	698	956	699
Control mean	0.00	0.00	0.06	0.06	0.02	0.04	0.01	0.02

Note: The main dependent variable is an index of women's report of sexual harassment from men from the same class as them (created using Anderson 2008 method). The questions asked female respondents about different types of sexual harassment perpetrated by men from their own class. The events are same as in the main results table on sexual harassment. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.16: Men’s beliefs on probability of women reporting to classmates and/or the complaints committee

What Would a Woman from Your Class Do in Hypothetical Sexual Harassment Situations?	Situation 1 (1)	Situation 2 (2)	Situation 3 (3)
<i>Panel A: She Will Report Formally</i>			
Male treatment	0.01 (0.02)	0.01 (0.01)	-0.01 (0.01)
N	1,310	1,310	1,310
Control mean	0.40	0.28	0.30
<i>Panel B: She Will Report Informally</i>			
Male treatment	0.05** (0.02)	0.04** (0.02)	0.03* (0.02)
N	1,310	1,310	1,310
Control mean	0.45	0.36	0.36

Note: The table shows regression results of estimating equations for the dependent variables on the class-level intervention for men. Number of observations is smaller here since only a sub sample of men was asked this question to keep survey duration in check. PDSLASSO is used for selecting controls in panel C. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table C.17: Men’s and women’s second order beliefs about their classmates

	Beliefs about male classmates			Beliefs about female classmates		
	Situation 1 (1)	Situation 2 (2)	Situation 3 (3)	Situation 1 (4)	Situation 2 (5)	Situation 3 (6)
<i>Panel A: Men’s Beliefs</i>						
Male treatment	0.07*** (0.02)	0.06*** (0.01)	0.04** (0.02)	0.07*** (0.03)	0.09*** (0.02)	0.06*** (0.02)
N	1,310	1,310	1,310	1,310	1,310	1,310
Control mean	0.52	0.35	0.38	0.56	0.38	0.43
<i>Panel B: Women’s beliefs</i>						
Male treatment	0.022 (0.024)	0.012 (0.019)	0.024 (0.019)	0.028 (0.026)	0.014 (0.025)	0.028 (0.024)
N	1,022	1,022	1,022	1,022	1,022	1,022
Control mean	0.59	0.33	0.36	0.64	0.43	0.45

Note: The table shows regression results of estimating equations for the dependent variables on the class-level intervention for men. The dependent variable is the percentage of other classmates who the men (women) think will agree with the law that the corresponding situation should be legally punishable in panel A (B). In columns 1–3, it shows second order beliefs about other men, and 4–6 shows that for about other women. Number of observations are smaller than the main tables since these were asked to only a sub-sample of students. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table C.18: Long-run effects of male intervention on perception of peers attitudes

	Perceived Social Support			Perceived defense			Perceived support			Perceived costs	
	Mild (1)	Inter (2)	Ext (3)	Mild (4)	Inter (5)	Ext (6)	Men (7)	Women (8)	via peers (9)	via ICC (10)	to networks (11)
<i>Panel A: All Men</i>											
Male treatment	0.07* (0.03)	0.04* (0.03)	0.05** (0.02)	0.03 (0.02)	0.01 (0.03)	-0.01 (0.04)	2.81* (1.61)	3.51** (1.63)	0.03* (0.02)	0.03* (0.02)	0.02 (0.02)
N	867	956	1048	879	765	468	1356	1352	1104	1079	1476
Control mean	0.66	0.80	0.85	0.87	0.74	0.64	69.8	66.3			
<i>Panel B: All women</i>											
Male treatment	-0.02 (0.04)	-0.06* (0.03)	0.01 (0.03)	-	-	-	1.14 (2.28)	1.29 (2.64)	-0.0141 (0.03)	-0.02 (0.02)	0.05 (0.04)
N	683	678	663	-	-	-	804	805	600	590	585
Control mean	0.62	0.68	0.77	-	-	-	55.78	64.3	0.87	0.87	0.35

Note: Perceived social support is a question on how likely the respondent thinks their classmates will punish someone engaging in sexual harassment of different types, perceived defense is how likely they think that their classmates will actively stop someone from harassing a female classmate, and perceived support from men and women is what fraction of their classmates will support a sexual harassment victim. Perceived costs beliefs about damage from being reported informally (column 9) versus formally (column 10), and costs to social networks for someone (column 10) who has a sexual harassment report against them.

Table C.19: Long-run effects of male intervention on self reported attitudes

	First order beliefs			Self reported attitudes		
	Mild	Inter	Ext	Empathy	Approach women	support victim
<i>Panel A: All Men</i>						
Male treatment	0.02 (0.02)	0.01 (0.02)	0.02* (0.01)	0.19** (0.07)	-0.55 (0.51)	0.03* (0.02)
N	923	979	1048	1096	1476	1147
Control mean	0.90	0.92	0.94	3.47	9.41	0.87
<i>Panel B: All women</i>						
Male treatment	-0.02 (0.03)	0.01 (0.02)	0.01 (0.03)	-	-	0.01 (0.02)
N	665	659	655	-	-	661
Control mean	0.85	0.85	0.89	-	-	0.90

Note: First order beliefs asks the respondent whether they would punish someone who sexually harasses a female classmate (mild, intermediate or extreme events), self reported attitudes highlight empathy in column 4, whether approached a woman to ask for date (5) and whether will support a victim of harassment (6). The PDSLASSO method is used to select controls, and strata and college fixed effects are included in the regressions. Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table C.20: Heterogeneity by year of study for opposite-sex relationships index

	Opposite- Sex Index (1)	Dating (Same Class) (2)	Opposite- Sex Friends (3)	Switches (Same Sex) (4)	Stays (Opp. Sex) (5)
<i>Panel A: All Women</i>					
Male treatment (β_1)	-0.134*** (0.042)	-0.017** (0.007)	-3.718 (2.99)	-0.126 (0.079)	-0.049 (0.074)
Older cohort (β_2)	0.190* (0.105)	0.007 (0.007)	6.510 (6.704)	0.012 (0.086)	0.197* (0.115)
Male treatment \times Older cohort (β_3)	-0.011 (0.060)	0.008 (0.013)	3.646 (4.174)	-0.026 (0.104)	0.064 (0.096)
$\beta_3 + \beta_2 = \beta_1$	0.005	0.059	0.130	0.481	0.063
N	1,381	1,144	1,354	555	525
Control mean	-	0.02	0.15	0.37	0.68

The table shows results from a regression of dependent variable for women on the class-level intervention for men. Older cohort is a dummy equal to 1 if the female respondent belonged to years two and three of the study and 0 otherwise. The dependent variable in column 1 is an index using Anderson (2008) created from a combination of dependent variables in columns 2–5. In column 2 the dependent variable is a dummy variable that asked men and women whether they were dating anyone in their own class or not, and in column 3 it is the proportion of opposite-sex friends from the same class reported by men and women. In columns 4 the dependent variable is whether the student switches from their same-sex partner from their own class or not, and in column 5 it is whether they stay with the opposite-sex partner from their own class or not. Note that there are a fewer number of observations for columns 4 and 5 because the lab in field was a between-subject design. Values are thus imputed using the KLK method for those who were not in a particular group. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. Randomization inference p-values are reported in square brackets. PDSLASSO is used for selecting controls. P-values adjusted for multiple hypothesis testing are reported as BH-adjusted p-values (Benjamini and Hochberg, 1995). Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table C.21: Treatment effect on men’s approach behavior toward women

Variable	Short Run (1)	Long Run (2)
Male treatment	-0.022 (0.036)	-0.029 (0.028)
N	589	582
Control mean	0.23	0.12

The dependent variable is a dummy equal to 1 if the woman was approached by a man in her own class and 0 otherwise. Strata fixed effects are included, and clustered standard errors are reported. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.22: Mediation analysis for short run

Sexual Harassment	Sexual Harassment from			
	Same-Class Index (1)	Mild Events (2)	Intermediate Events (3)	Extreme Events (4)
<i>Panel A: No Controls, All Women</i>				
Male treatment	-0.044 (0.027)	-0.005 (0.017)	0.013 (0.009)	-0.009** (0.004)
Dating in same class	0.375 (0.266)	0.091 (0.089)	0.018 (0.052)	0.049 (0.047)

Note: The table reports results from a regression of the dependent variable on the class-level male intervention dummy variable and on an indicator for whether a female student reports dating a male student from the same class or not. Sexual harassment variables are the same as before. The dependent variable in column 1 is an index created using the Anderson (2008) method. This index combines questions on different types of sexual harassment perpetrated by men in the same class, as reported by women in columns 2–4. Clustered standard errors are in parentheses, and strata fixed effects are included in all specifications. PDSLASSO is used for selecting controls. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table C.23: Sexual harassment for single women in the control group

	N	Mean	SD	Min	Max
Mild events	494	0.07	0.25	0	1
Intermediate events	488	0.03	0.17	0	1
Extreme events	488	0.01	0.09	0	1

Note: The table shows the prevalence of sexual harassment among women in the control group who report being single in the endline survey.

Table C.24: Spillovers on men in control classes

Dependent Variable	Overall Knowledge (1)	Intrinsic Attitudes		Perceived costs	
		Victim-Blaming Attitudes (2)	Legal Costs (3)	Perceived Social Costs (4)	
<i>Panel A: Male friends outside</i>					
N male friends from outside own class	-0.0421 (0.012)***	.017 (0.048)	0.0089 (0.017)	-0.0034 (0.022)	
N male friends from treatment classes outside own class	0.0001 (0.025)	0.028 (0.112)	0.0277 (0.032)	-0.0179 (0.048)	
N	697	678	697	687	
<i>Panel A: Female friends outside</i>					
N female friends from outside own class	0.029 (0.029)	0.115 (0.092)	0.002 (0.037)	-0.033 (0.044)	
N female friends from treatment classes outside own class	-0.039 (0.051)	-0.196 (0.213)	-0.091 (0.059)	-0.006 (0.068)	
N	697	678	697	687	
Mean	0.00	2.15	0.00	0.00	

Note: The sample consists of only men from control group. The dependent variables are same as for table (add XXXX number). Asterisks denote significance: *p<0.1, **p<0.05, and ***p<0.01.

Table C.25: Short and Long run effects of the Female intervention on mechanism outcomes for Women

	Awareness		Perc Social Costs		Perc legal Costs		Intrinsic Attitudes*	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
<i>Panel A: All women</i>								
Female treatment β_1	-0.01 (0.05)	-0.07 (0.09)	0.02 (0.06)	-0.07 (0.11)	0.08 (0.09)	-0.02 (0.12)	0.03 (0.07)	0.13 (0.29)
High intensity β_2	-0.02 (0.07)	0.13** (0.07)	-0.06 (0.08)	-0.11 (0.12)	0.10 (0.10)	-0.16 (0.12)	0.03 (0.23)	0.35 (0.23)
High intensity \times Female treatment β_3	0.03 (0.07)	-0.18 (0.12)	-0.05 (0.10)	0.06 (0.18)	-0.18 (0.12)	0.02 (0.17)	-0.18 (0.31)	-0.82** (0.41)
N	637	359	637	368	637	368	629	368
$\beta_2 + \beta_3 - \beta_1 = 0$	0.93	0.57	0.28	0.86	0.62	0.62	-	-
Control mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: The table reports results from a fixed effects regression of the dependent variable on the treatment intensity dummy, individual female treatment and the interaction of the two using short and long-run surveys with the same sample of students. The dependent variable in column 1 is an index created by combining questions used to test awareness of sexual harassment. In column 2 the dependent variable is an index created using questions on the perception of social disapproval from other classmates, and in column 3 it is an index created using questions on the perception of legal support for victims. In column 4 the dependent variable is the coefficient on the interaction between the treatment and list treatment; In column 5, the dependent variable measures donations to an NGO that works against sexual harassment. The PDSLASSO method is used to select controls, and strata fixed effects are included in the regressions. Asterisks denote significance: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.