Debiasing Law Enforcement Officers: Evidence from an Expressive Arts Intervention in India*

Sofia Amaral† Girija Borker‡ Nishith Prakash§ Maria Micaela Sviatschi¶

March 26, 2026

Preliminary - do not cite without permission.

Abstract

Biases among state agents can undermine the quality of public service delivery and restrict access to justice, particularly for marginalized groups. We study whether reshaping police officers' attitudes toward gender-based violence (GBV) improves real-world policing performance. We partner with the Government of Bihar, India, to implement a clustered randomized controlled trial covering 419 police stations and approximately 2,500 senior officers serving 42 million citizens. The novel intervention builds on an expressive arts–based pedagogy that uses role play, perspective-taking, and experiential learning to address gender norms, empathy, and technical skills related to GBV. Six months after the intervention, treated officers exhibit significant improvements in attitudes toward GBV, reduced victim-blaming, higher empathy, and greater legal knowledge. These improvements translate into meaningful changes in workplace environment and job performance. Eight months after the intervention, junior female officers in treated stations report significantly lower levels of workplace anxiety and harassment. One year later, standardized victim audits reveal substantial improvements in officers' responses to women seeking help, including lower rates of case dismissal and more appropriate guidance. Exploring the random allocation of officer transfers, we show that even when treated officers move, individual attitudes toward GBV improve in relation to their control peers. We also find evidence of positive spillovers among officers who are only exposed to the treatment via interaction with treated officers. Together, these findings demonstrate that implicit biases among frontline state agents are malleable and that immersive behavioral training can generate persistent improvements in institutional performance and workplace culture.

Keywords: Gender-based violence, policing, workplace well-being, pedagogy of the oppressed, experiential learning, social norms, spillovers.

*We thank KS Anupam, Chandra Shekhar Kumar, Vinay Kumar, Rajyabardhan Sharma, Amir Subhani, Praveen Vashistha, Rahul Singh, Anil Kishore Yadav, GS Gangwar, and Vikas Vaibhav. We are also thankful to the following Superintendents of Police (SPs): Har Kishore Rai, Kantesh Mishra, Manavjit Singh Dhillon, Ashok Mishra, Rakesh Kumar, Pramod Kumar Yadav, Swarn Prabhat, Shailesh Kumar Sinha. This work was supported by the ifo Institute, Fund for Innovation in Development (FID), J-PAL Crime and Violence Initiative, J-PAL Gender Initiative, The World Bank, and Princeton University. The views expressed here do not necessarily represent the views of our partners. Ethics approval was obtained from Princeton University.

†South Asia Gender Innovation Lab, World Bank. Email: s.amaral@worldbank.org

‡DIME, World Bank. Email: gborker@worldbank.org

§Northeastern University, CESifo, IZA, CReAM, HiCN, & GLO. Email: n.prakash@northeastern.edu

¶Princeton University. Email: msviatschi@princeton.edu

1 Introduction

The quality of public service delivery in developing countries remains persistently low despite it being an necessary condition to economic development (Besley and Persson, 2009; Page and Pande, 2018). By limiting citizens' access to essential services, poor policy delivery undermines state capacity. A large literature documents weak performance across sectors such as education and health, often driven by inadequate training, misaligned incentives, and weak accountability (Finan et al., 2017; Das et al., 2016). These challenges are particularly acute in law enforcement. Unlike many other public services, policing has no close private substitute: the state retains a monopoly over the application of the law. As a result, poor police practices—especially those shaped by bias—can severely undermine public trust, deter service demand, and restrict access to justice. These concerns are especially salient for gender-based violence (GBV), where dismissive responses, victim-blaming, and procedural barriers at police stations reduce reporting and perpetuate inequality in legal protection.[1]

In response to these challenges, policymakers have turned to gendered and GBV segregated polices and gender quotas, which have increased reporting in some settings but face important limitations in scale and implementation (Miller and Segal, 2019; Sukhtankar et al., 2022; Sviatschi and Trako, 2024; García-Hombrados et al., 2024).[2] In spite of these efforts, policing remains a male-dominated and highly hierarchical profession, where senior officers exercise substantial discretion and tenure protections constrain personnel reallocation. These institutional features raise a central question: how can the quality of the existing police force be improved when bias among frontline state agents limits performance?

This paper studies whether reshaping the attitudes and skills of senior police officers

[1] Despite high prevalence of GBV, women's reporting rates remain extremely low, reflecting both social stigma and institutional failures. Prior work shows that police officers' attitudes toward GBV—such as beliefs about victim credibility or whether sexual harassment merits police action—strongly shape officers' behavior and affect crime incidence (Amaral et al., 2025). Biases within police forces therefore constrain the effectiveness of law enforcement responses to gender crimes and erode fairness and public trust.

[2] In India, some states have committed to gender quotas in law enforcement. Yet, recruitment of female officers remains low and in many instances the minimum threshold of quotas is not met (Tata Trusts, 2023).

through a novel expressive arts training can improve policing performance and the workplace environment in the presence of entrenched gender bias. To study this question, we partner with the Chief Minister of Bihar and the Director General of Police in India, to implement a clustered randomized controlled trial covering 419 police stations across twelve districts and over 2,000 senior officers serving 42 million citizens. Bihar is the poorest state in India, and provides a particularly relevant context: despite recent legislative improvements, gender violence remains highly prevalent, reporting rates are extremely low, and women remain underrepresented in decision-making roles within police stations.[3] Our intervention targets key decision-makers—Station House Officers, Inspectors, Sub-Inspectors, and Assistant Sub-Inspectors—who shape both citizen-facing behavior and internal workplace norms.[4]

Motivated by evidence that immersive and experiential learning can be more effective than traditional instruction in shifting beliefs, norms and decision-making in high-discretion settings (Bursztyn and Jensen, 2017; Andries et al., 2024; Bursztyn et al., 2020; Dube et al., 2025; Ashraf et al., 2024), we design a novel expressive arts–based training tailored to the institutional realities. In particular, building on extensive qualitative fieldwork—including sustained observation in police stations and interviews with officers, survivors of gender-based violence, and civil society organizations—we document that policing organizations are highly hierarchical and characterized by deeply entrenched gender norms, with senior officers exercising substantial discretion over both case registration and workplace culture. These insights motivated an intervention targeting officers at the top of the hierarchy and emphasizing perspective-taking and experiential learning.

The intervention draws on the Freirean principle of conscientization (Freire, 1970), which posits that social liberation requires the oppressed to develop a critical awareness

[3] According to the Bihar Economic Survey 2025-26, Bihar's per capita remains the lowest in India with a per capita GDP at current prices of USD 925. Bihar's GDP is comparable to those of other low-income countries such as Sierra Leone, Uganda, or Gambia.

[4] Reforming the police has been a significant challenge in Bihar. Despite recent legislative improvements, GBV remains a major epidemic in the state for which concrete prevention measures have not yet been taken. For example, the latest estimates put Bihar as one of the worst performing states regarding the incidence of intimate-partner violence. At the same time, despite the existence of gender quotas in police recruitment the presence of women in police stations in decision-making roles is minor.

of their structural reality. This pedagogical approach was translated into the practice by Boal (2014), whose *Theatre of the Oppressed* methodology replaces the passive 'spectator' with the 'spect-actor'—an individual who rehearses social change on stage to prepare for real-world action. The intervention method was adapted to the context of the dynamics of the police and its interactions with female GBV survivors. It included role play, perspective-taking exercises and simulated police station scenarios, to immerse officers in real-world interactions with survivors of GBV to build empathy and strengthen officers' technical and legal knowledge. Through structured perspective-taking, officers are encouraged to critically examine the viewpoints of all parties involved, assume the roles of victims and responding officers, and collectively reflect on appropriate conduct. These experiential components are complemented by guided reflection, peer feedback, and explicit commitments to station-level practice changes, creating repeated behavioral feedback loops that link awareness to practice. Unlike traditional classroom-based trainings, the program is designed to address implicit biases within the police hierarchy while simultaneously strengthening officers' technical and legal skills.

Our analysis draws on a rich set of newly collected data that captures officer attitudes, behavior, and workplace dynamics using surveys, random victims' audit data, and administrative police records. We measure police officers' mindsets, attitudes, and technical skills related to GBV using detailed baseline and endline surveys administered to senior male officers. We make use of self-reported measures, along with vignettes, thus, taking into account possible measurement concerns (Bursztyn et al., 2025). To assess changes in workplace environment, we conduct confidential surveys with junior female constables stationed at treated and control police stations, collecting information on harassment, anxiety, and perceptions of senior officers. To measure real-world policing behavior, we implement a standardized victim audit in which trained decoys visit police stations to lodge complaints and systematically record officers' responses, and we complement these measures with administrative data on crime registration and handling drawn from approximately 600,000 official police records. Together, these data allow us to link changes in officer attitudes to downstream effects on workplace culture and service delivery.

We find that immersive expressive-arts training meaningfully improves police per-

formance, workplace environment, and interactions with survivors of GBV. Four to six months after the intervention, treated senior officers exhibit significant improvements in attitudes toward GBV, including higher empathy, reduced victim-blaming, and stronger legal and procedural knowledge. These changes translate into improvements in workplace culture: eight months after the training, junior female officers stationed with treated seniors report lower levels of harassment and anxiety. One year after the intervention, standardized victim audits reveal substantial improvements in officers' real-world behavior toward women seeking help, including lower rates of case dismissal, reduced victim-blaming, and more appropriate guidance and referrals. Together, these results demonstrate that implicit biases among frontline state agents are malleable and that immersive behavioral training can generate persistent improvements in institutional performance and workplace culture.

We also provide evidence that the intervention operates primarily by shifting the attitudes and practices of officers for whom bias was most constraining at baseline, rather than reinforcing already progressive beliefs. In particular, we find that treatment effects are concentrated among officers who exhibited lower baseline levels of gender-sensitive attitudes toward GBV. Heterogeneity analyses show that officers with weaker initial attitudes experience significantly larger improvements in overall GBV skills, empathy, and reductions in victim-blaming following the training, while officers with more progressive baseline views exhibit smaller and statistically insignificant changes. These findings are consistent with the training addressing implicit biases and entrenched norms, and they indicate substantial scope for improving policing performance by targeting officers with the greatest initial attitudinal gaps.

To further shed light on the mechanisms underlying the intervention's effects, we exploit variation generated by officer transfers across stations.[5] The transfer-based heterogeneity results indicate that treated officers who remain in treated stations experience significant improvements across skills, attitudes, and technical competencies, while those

[5]Officer transfers are an important feature of the context of policing in India and are in large part determined by general law enforcement necessities. This feature has been explored by Iyer and Mani (2012) in the context of bureaucrats. We show transfers are exogenous to treatment assignment.

who transfer to control stations continue to display gains in key domains, suggesting that training induces durable individual-level changes consistent with a persistent change in attitudes towards gender violence. At the same time, control officers who move into treated stations exhibit sizable improvements, particularly in skills and technical knowledge, pointing to the importance of peer exposure and social-norm formation. The absence of significant differences between treated stayers and treated movers further indicates that treatment effects travel with individuals, whereas improvements among control movers highlight environment-driven spillovers. Taken together, these patterns suggest that the intervention operates through a combination of internalized attitude change, norm diffusion within stations, and enhanced procedural knowledge, with transfers providing a useful empirical lever to distinguish between individual persistence and contextual reinforcement.

Several pieces of evidence suggest that our results are unlikely to be driven by alternative mechanisms such as changes in incentives, compositional effects, social desirability bias, or increased monitoring. The intervention did not alter officers' formal responsibilities, compensation, promotion prospects, or enforcement intensity, ruling out incentive-based explanations. Results are robust to controlling for officer transfers and station-level staffing changes, and we do not observe differential attrition across treatment and control stations, mitigating concerns about compositional shifts. We also show that results are not driven by changes in social desirability bias, or driven by officers with a larger degree of desirability bias.[6] The fact that we also detect effects in audit data, and among non-treated officers (females and non-treated transferred officers) further reassures that results are not driven by this measurement concerns. Next, we also find that treatment effects do not vary with officers' rank or hierarchical position, which would be expected if changes were primarily driven by senior officers updating beliefs about peers' norms and rewarding visible compliance. Instead, improvements are concentrated among officers with weaker baseline attitudes, consistent with internalized changes in attitudes and

[6] As we will explain, our novel measures include a mix of direct-questioning and vignettes – tools designed to mitigate concerns over desirability bias (Bursztyn et al., 2025).

decision rules rather than social image management or perception updating alone. Moreover, outcomes are measured well after the training—up to one year later—and in settings where officers were unaware of being evaluated, such as standardized victim audits, reducing the scope for short-run signaling or demand effects. Finally, treatment effects do not vary with the share of female constables or observable station characteristics, suggesting that results are not driven by contemporaneous changes in station composition or peer exposure. Together, these patterns support an interpretation in which the intervention operates through shifts in officer attitudes, norms, and decision-making rather than through mechanical or institutional channels.

This paper contributes to the literature on public service delivery (Finan et al., 2017) by providing causal evidence that implicit biases among state agents constrain performance in high-discretion institutional settings and can be durably reshaped. Prior work shows that beliefs, attitudes and social norms affect decision-making in courts and policing (Amaral et al., 2025; Dube et al., 2025; García-Hombrados et al., 2024; Harris et al., 2024), but there is limited evidence on whether interventions can generate persistent behavioral change among frontline bureaucrats or service providers – key agents in GBV response. We show that an immersive, experiential training targeted at senior police officers significantly improves both attitudes and real-world behavior toward victims of gender-based violence, establishing a direct link between bias reduction and service quality in a core state function.

Our findings add to the work of (Dube et al., 2025; Harris et al., 2022). Dube et al. (2025) shows that experiential training can similarly induce a 'cognitive pause' in high-stress, high-discretion policing environments and lower police use of force. We add to this by showing that these cognitive shifts can be leveraged to durably improve the quality of service delivery and survivor-centric engagement in GBV cases by pinpointing the behavioral and normative channels that promote such changes in officers mindset. Harris et al. (2024) finds that identity-based ethics training for police in Ghana improves professional values and reduces self-reported willingness to engage in unethical behavior. We extend this evidence by showing that such behavioral shifts can extend to the empathetic and technical quality of service provided to victims of GBV.

Conceptually, our findings build on and extend behavioral interventions (Andries et al., 2024; Alan et al., 2021; Canales et al., 2025), which demonstrates that perspective-taking and empathy can shift beliefs and behavior by exposing individuals to others' lived experiences. While previous literature on the intersection of gender and male's social norms focuses on adolescents and students (Dhar et al., 2022; Sharma, 2022; Matavelli, 2024; Brar et al., 2025) and intra-household dynamics (Bursztyn et al., 2020; Cunha et al., 2023), our study applies perspective-taking to a novel context – institutional decision-makers operating within rigid hierarchies and under high discretion. We show that immersive role reversal and experiential learning—embedded within officers' professional environment and reinforced through peer interaction and practice—lead not only to belief changes but also to internalized shifts in attitudes/norms and decision rules that persist in private, high-stakes interactions.[7]

Our second contribution is to the literature on the scalability and social change and norms interventions. While previous work has primarily analyzed the direct effects of individual-level interventions Dube et al. (2025); Harris et al. (2022); Canales et al. (2025), we provide novel evidence on the differential spillover effects of GBV interventions within hierarchical organizations. We document a bifurcation in how skills diffuse: while soft skills—such as empathy and bias-mitigation—manifest positive spillovers among untreated peers, technical competencies remain strictly localized to the trained agents. This suggests that the 'social' aspects of GBV training are susceptible to diffusion via peer contagion, whereas the 'technical' aspects are more likely to be constrained by institutional rigidities and the lack of systemic standardization.

Finally, this study contributes to a novel strand of the literature seeking to understand the impacts of workplace dynamics (Folke and Rickne, 2022; Adams-Prassl et al., 2024; Dahl and Knepper, 2021; Boudreau et al., 2023). Previous work shows that hierarchy and gender are important drivers of under-reporting (Adams-Prassl et al., 2024), and that

[7] A recent systematic review on gender norms interventions concludes that there is a lack of studies targeted to understand males, non-student populations, and that go beyond short-term effects and self-reported measures (Stewart et al., 2021). Our study fills this gap on all dimensions of scientific gaps identified.

improvements in outside options can improve reporting rates (Dahl and Knepper, 2021). We add to this literature by showing that improvements of social norms among men in high hierarchy positions have positive effects on their female subordinates.

2 Background

In order to unpack the complexity of policing toward GBV survivors in our context, we completed qualitative data analysis, in-station observation exercises, and various consultations with NGOs and police stakeholders. This informative work allowed the development of the theory of change, quantitative survey tools, and the intervention. In this section, we provide a description of the context and response system of the police.

2.1 GBV and the Police in Bihar

Gender-based violence continues to be pervasive in Bihar. More than 20,000 cases of violence against women were registered in 2022 in the state (National Crime Records Bureau, 2023). This figure saw an increase of 13% in 2022 as compared to the previous year.[8] Findings from Indian Institute of Population Sciences (2022) also show that 40% of women aged between 18-49 years have experienced spousal physical or sexual violence from their current partner. Further, only 11% of women who reported facing physical violence sought some form of assistance, while 7% never sought any assistance, but did confide in someone.Data from the DHS also shows that only 2% of women who faced physical violence approached the police (Indian Institute of Population Sciences, 2021). The lack of demand may be due to the stigma surrounding gender violence in Bihar wherein the police reportedly state that cases are false (Datta and Satija, 2020).

In the last decade, there has been a significant push by the State Government of Bihar towards increasing the representation of women in the police. In a landmark policy decision, the Chief Minister of Bihar, as of 2013, announced a 35 per cent reservation for

[8]The media article can be read here.

women in police recruitment from the rank of constables to Deputy Superintendent of Police.[9] As of 2020, Bihar surpassed Himachal Pradesh (19.15 per cent) and Tamil Nadu (18.5 per cent) to become the state employing the highest share of female police, with around 25.3 per cent of the total strength, which is a stellar improvement from 2015 where females accounted for merely 3.3 per cent of the police force.

While remarkable progress has been made to improve the representation of women in Bihar Police, it is important to note that majority of this representation is only confined to the lower ranks of this hierarchical set-up. For instance, only 2 out of 11 Director General (DG) rank, merely 3 out of 29 Additional Director General (ADG) and 6 out of 33 Deputy Inspector General (DIG) rank IPS officers are female, while none of the 17 Inspector General (I.G.) rank IPS officers in the state are female. Also, there are only 5 out of 88 female Superintendent of Police (SP) as of March, 2023.[10] As a result, women's presence in the police remains at the Constable level. According to our estimates, there are only 6 out of 419 (1.43%) female SHOs, 0 female inspectors, 230 out of 1779 (12.9%) female sub-inspectors, 39 out of 1618 (2.3%) female additional sub-inspectors. However, in the constable rank women are 27.3% of the police strength – see Table A.1.

In addition, the police system in Bihar is organized hierarchically, with senior officers making the key decisions. It is headed by the Director General of Police (DGP), who is the highest-ranking officer in the force. At the district level, there are the Superintendent of Police (SP)[11] and Deputy Superintendent of Police (DySP). The senior officers are responsible for making decisions on matters such as deployment of personnel, allocation of resources, and implementation of policies.

Within a police station, the Station House Officer (SHO), who is usually an Inspector or, in some cases, a Sub-Inspector (SI) rank officer, is the officer in charge of the station. Working under the close supervision of the SHO are the Sub-Inspector (SI) and Assistant

[9] Refer to this link

[10] Accessed from Home Department of Bihar, as of 01 March 2023. https://home.bihar.gov.in/CMS/IpsView.aspx

[11] In some large districts such as Patna and Muzaffarpur, we instead have Senior Superintendent of Police (SSP) as the highest-ranking officer at the district level.

Sub-Inspector (ASI) rank police officers, who are responsible for registering FIRs, conducting investigations, collecting evidence, and making arrests.

Lastly, there are several lower-ranked police officers known as Police Constables, who work under the close supervision of the other higher-ranked officers within a police station and assist them in their duties. Constables typically do not hold decision-making roles, but they are often the first point of contact with citizens, and their experiences during these initial stages are taken into account by senior officers when making decisions and documenting case procedures.

Guided by extensive qualitative fieldwork conducted over 1.5 years—including sustained observation in police stations and interviews with officers, survivors of GBV, and civil society organizations—we identified three institutional constraints that limit access to justice to GBV survivors. First, police stations operate within rigid hierarchies where senior officers exercise substantial discretion over case registration and workplace norms. Second, gender biases are deeply embedded in everyday practices, often manifesting as victim-blaming, skepticism toward complaints filed by women, and reliance on informal compromise rather than formal legal action. Third, traditional classroom-based training was widely perceived as ineffective and rarely translated into behavioral change.

These insights motivated the design of a novel, immersive training targeted at senior officers. Rather than attempting to change the composition of the police force—through quotas or specialized women's units—we focused on reshaping the attitudes and skills of existing decision-makers within the hierarchy.

The intervention targeted active-duty senior officers—Station House Officers (SHOs), Inspectors, Sub-Inspectors (SIs), and Assistant Sub-Inspectors (ASIs)—currently posted in the field. These officers are the first point of institutional contact for survivors and are responsible for registering First Information Reports (FIRs), initiating investigations, and supervising constables. In practice, their beliefs shape both citizen-facing behavior and internal workplace culture.

Training was delivered in batches of approximately 15–30 officers drawn from multiple police stations within the same district. To preserve operational capacity, only a subset of

senior officers from each station participated at any given time, and only one SHO per station was invited. Implementation was coordinated district-by-district in collaboration with designated nodal officers, ensuring institutional endorsement while maintaining the integrity of police operations.

2.2 Expressive Arts Training and the Theatre of the Oppressed Method

The intervention curriculum was developed in collaboration with theater practitioners and gender experts and drew on principles from the theater of the oppressed technique (Freire, 2020) and other applied theater techniques (Boal, 2014). The central premise was that experiential learning and perspective-taking are more effective at reshaping implicit biases and social norms than lecture-based instruction. The principles of immersive and experiential learning techniques have been found to be effective in different contexts in various studies (Canales et al., 2025; Ashraf et al., 2024; Dube et al., 2024; Andries et al., 2024).

To design the curriculum, we collaborated with Mittika, a Delhi-based NGO, and got input from several high-ranked Indian Police Services (IPS) Officers, lawyers, and civil society members.[12] The research team was extensively involved in the development of the training curriculum to ensure that the module was closely aligned with the theory of change.

The training lasted three consecutive days and followed a structured progression. The first phase focused on making implicit gender stereotypes explicit. Officers were asked to reflect on their vision of an "ideal police station," their professional responsibilities, and the gap between institutional ideals and everyday practice. Interactive exercises prompted participants to categorize traits and roles as "male" or "female," followed by collective reflection on how these associations are socially constructed rather than biologically determined. The objective was not to prescribe correct answers, but to create space

[12]Mittika is a major NGO developing gender training programs in India, Afghanistan, and Tibet. It has several years of experience working with diverse groups of people ranging from adolescents to public officials.

for introspection and peer discussion.

The second phase introduced structured discussions on inequality, access to justice, and the lifecycle of violence. Facilitators enacted realistic police station scenarios in which officers mishandled GBV complaints—for example, dismissing a rape survivor or suggesting informal compromise. Participants were invited to interrupt the scene, identify procedural errors, and reenact the interaction according to correct legal and professional standards. In subsequent exercises, officers alternated between playing survivors and responding officers, while peers observed and provided feedback. These role-playing sessions required participants to engage in deliberate perspective-taking—articulating the experience of a complainant within her social context—and to rehearse alternative responses.

The final phase linked attitudinal change to concrete institutional practice. Officers reviewed relevant GBV legislation and procedural requirements, participated in quizzes reinforcing legal knowledge, and received a practical handbook summarizing key provisions. Sessions emphasized survivor-centric policing, including private intake, non-judgmental language, and appropriate referral pathways. Each participant concluded the training by formulating specific action points to implement in their station.

Throughout the program, interactive games reinforced learning. For example, in a "Snakes and Ladders" exercise, officers advanced or regressed based on responses to hypothetical complaint-handling scenarios, illustrating how procedural shortcuts can harm victims and erode institutional credibility. In a "Web of Relations" activity, participants mapped interdependencies within the police station, highlighting how individual actions shape broader workplace culture. For further details on the intervention see the Online Appendix.

The intervention was designed to operate through three channels. First, an attitudes/preference channel. Perspective-taking and role-play exercises were intended to increase empathy and shift other-regarding preferences. In particular, the training sought to increase the weight officers place on complainants' welfare when exercising discretion in gender-based violence cases. If this channel dominates, we should observe improvements in empathy

and reductions in victim-blaming even in settings where peer exposure is limited, and effects should be strongest among officers with initially regressive attitudes. Moreover, changes should appear not only in peer-facing contexts but also in individual-level measures of perspective-taking.

Second, a new social norm channel. Because the training was conducted collectively and publicly among senior officers, it was designed to create a new norm among seniors on what is expected within the station. In hierarchical organizations, behavior is shaped not only by private attitudes but also by beliefs about peer norms and supervisory expectations. If this mechanism is operative, behavioral changes should extend beyond individual attitudes to workplace culture and interactions with junior officers. This channel can be distinguished from pure preference change by examining spillovers to female constables' reported workplace experiences and to non-treated officers.

Importantly, the collective and performative nature of the training may also enhance recall through social learning. Officers did not passively receive information; they enacted scenarios, interrupted scenes, defended decisions, and publicly negotiated appropriate conduct. Learning that is enacted and socially observed generates shared episodic memory and repeated peer cues in daily work environments. If recall and norm reinforcement play a role, we would expect persistence of treatment effects over time and limited decay as days since training increase, consistent with a self-reinforcing equilibrium within stations.

Third, a technical-knowledge channel. The curriculum included structured instruction on GBV-related procedures and legal provisions. At baseline, officers with more regressive attitudes also displayed weaker legal understanding, suggesting that attitudinal bias and procedural gaps were intertwined. By pairing reflection with procedural reinforcement, the intervention sought to ensure that any shift in attitudes translated into correct implementation. This mechanism can be empirically separated from the preference and norm channels by examining improvements in objective legal knowledge measures and protocol-based outcomes that do not rely on subjective reporting.

Finally, because the intervention targeted senior officers — the key decision-makers

within police stations — any changes in preferences, new norms, or procedural knowledge were expected to diffuse downward through supervision and daily interactions. Evidence of spillovers to junior female constables and to real-world complainant interactions provides a way to assess whether behavioral change reflects an individual-level shift or a broader institutional adjustment.

Unlike reforms that alter staffing composition or create parallel institutional structures, this intervention seeks to improve the performance of the existing state apparatus by reshaping the beliefs, expectations, and skills of incumbent decision-makers. By distinguishing preference change, norm updating, and technical reinforcement, the design allows us to test whether bias within state agents operates through malleable channels — and whether experiential, collective training can translate into persistent changes in institutional behavior and access to justice.

2.3 Implementation

The training of trainers (ToT) was organized in September 2022, which included a six-day intensive residential training and dry runs of the Police Training Curriculum – refer to Appendix-B. The trainers were trained by Mittika and Jaya Iyer – a lead expert in the use of applied theater techniques, and Ms. Sutapa Sanyal (IPS), a retired Director General of Police. The facilitators were consistently graded by the lead trainers on several parameters such as understanding of GBV and related issues, content delivery, energy and conviction, teamwork, and process flow. At the conclusion of the ToT and pilot, 14 out of 19 trainers, i.e., about 74% of the facilitators, were retained for the main intervention. All trainers responded to a short survey to measure aspects related to the intervention, their experience, and other socioeconomic outcomes.

The intervention was implemented in a staggered manner across districts in order to accommodate any possible law and order situations or festivities when the police are extremely busy. For every district, we sought approvals from the SP before the implementation of any survey or inviting officers for the training intervention. Afterward, for every activity, permissions were requested ahead of surveys and invitations for the training. The

target batch size for implementing the 3-day curriculum was set at 25, and officers were invited from different stations at different points in time in order to avoid disrupting the daily activities of the police station. Overall, training take-up was on average 84%, and attendance was on average 2.61 sessions. Police officers who attended the intervention received approximately 35 hours of exposure to the curriculum.

3 Data

Officers survey. The baseline officers' survey was conducted in a staggered manner in our sample districts, following a pilot survey conducted in Arwal in September 2022.[13] Participation in this survey was motivated as part of a professional task that the officers had to complete. Given that we had received permission from the respective SPs of the districts, the officers were comfortable sharing their responses with our survey team. The survey module was comprehensive, and it took, on average, 45 minutes to administer.

The survey instrument collected: (i) demographic and socio-economic characteristics of the officers, (ii) novel measures of victim-blaming, (iii) truthfulness perception of GBV reports, (iv) acceptance towards gender and GBV, (v) empathy, (vi) a novel scale to measure behaviors associated with externalization of responsibilities by the officers, (vii) technical and legal knowledge of the officers regarding GBV, well-being of officers (GAD and PHQ score), and a personal rigidity scale (Rehfisch, 1958).

We also measure social desirability bias based on Crowne and Marlowe (1960). Finally, the instrument also included a time-use survey within the workplace. We compute indices from the different scales. The indices are described in Appendix B.

The officers' endline survey was conducted in a staggered manner, 3-5 months since the conclusion of the intervention in a district. The endline survey collected the same outcome measures as the baseline.

[13] Arwal is a non-sample district that was chosen as a pilot district given its proximity to the base location of our study, i.e., Patna, and the fact that it is a relatively small district with a total of 12 police stations and outposts as of September 2022.

Female constables' survey. We hypothesize that the intervention creates a change in the behaviors and attitudes of trained male officers towards the females in their workplace. To test this hypothesis, we conduct a survey of the female constables in the sample police stations. The survey includes information on demographic and socio-economic characteristics, perceptions of workplace roles of female constables about male and female colleagues of all ranks. We also collect information on workplace empowerment, voice in the workplace, and harassment within the workplace. This survey is administered by female enumerators only.

Decoy. The decoy exercise was conducted in 6 districts in 185 police stations out of the total of 419 police stations, i.e., nearly 44% of our sample. The decoy visits took place in both treatment and control police stations. The exercise closely followed the approach used by Banerjee et al. (2021), wherein a decoy surveyor was sent to all the police stations of the decoy study districts to enact a case and to observe the response of the police officers, particularly if they were willing to register the case. Per station, trained actors attempted to file a complaint of gender-based violence and of a non-gender-based violence offense. Stations selected to participate in the decoy experiment are comparable to those that were not selected. Each police station was surveyed three times over a pre-agreed time period (decided in conjunction with the district authorities). The timing of decoy visits was randomized across waves to mitigate communication among police stations. Overall, we completed 185×3=555 decoy visits.

Administrative police records. We collected all FIRs registered at sample police stations from the State Crime Records Bureau. The data contains around 600,000 complaints.

3.1 Additional Data

Enumerators survey: To study potential enumerator effects, we conducted an enumerator survey before the start of the baseline survey in September 2022. The survey was administered after the enumerators were selected to be part of the survey team. It was self-completed on tablets and was voluntary. We surveyed the enumerators' team on socio-economic characteristics, including age, education, birthplace, number and gender

of children, life satisfaction, gender norms, and personality traits following Di Maio and Fiala (2020).

Reflection survey: To understand immediate reactions of participants to the training, we collected a reflection survey in both treatment and control stations. We use this information to evaluate the differential effects on officers' empathetic concern and perspective taking. In the treatment group, the surveys were conducted in the respective training venues upon the conclusion of the final day of training. Additionally, a set of training feedback-related questions were also asked in the reflection survey for the treatment group in order to capture an immediate self-reported measure of satisfaction with the training curriculum. The reflection surveys for the control group were administered at the police stations, and our team of enumerators visited the individual stations during the same month in which the training sessions were conducted for the respective districts.

Facilitators survey: To study potential effects of the individual differences in the characteristics of the facilitators, we conducted a facilitator survey before the start of our pilot training in September 2022. The survey was administered after the facilitators were selected to be part of the training team, following the Training of Trainers (ToT), which was conducted between September 12 and 17, 2022, and Pilot Training Sessions in Arwal, which were conducted between September 18 and 21, 2022. Participation in this survey was voluntary, and the surveys were self-completed on tablets or smartphones via a link that was sent to the facilitators.

We surveyed the team of facilitators on socio-economic characteristics, including age, education, birthplace, marital status, and number of children. We also asked specific questions related to their work experience and familiarity with the context of working with the Bihar Police. The survey instrument also had questions to capture the perceptions of the facilitators on the current state of policing towards gender-based violence crimes. Finally, the instrument included questions to capture the individual differences in personality traits.

Police station facilities survey: For balancing, we collect information on station infrastructure quality and calculate a police station facility index. The index is a simple count of

services available at stations. Through the police station facilities survey, we also measure the extent to which the infrastructure of stations is friendly to female officers.

4 Experimental Design

This section describes the randomized experimental design, data collection procedures, and study protocols. We outline the empirical strategy and present evidence on randomization, balance, and compliance.

4.1 Intervention and Design

Our study takes place in 12 districts in Bihar: Bagaha, Bettiah, Bhojpur, Gopalganj, Motihari, Muzaffarpur, Nalanda, Patna, Saran, Sitamarhi, Siwan, and Vaishali (see Figure 1). This area was selected based on consultations with the Additional Director General of Police (Weaker Section) and the Special Secretary to the Home Minister, as well as the number of stations and police officers required for the study.

We completed the baseline for 3,526 police officers across 419 police stations in the 12 districts. The study was implemented across these 12 police districts and includes a total of 419 police stations and outposts.[14] Officers sample constitutes the sample of senior officers posted in the sample stations. The state of Bihar has a total of 1,094 police stations and outposts (National Herald India, 2022)[15], so our sample covers nearly 37% of the total police stations in the state.

We use a clustered randomized control trial experiment to identify the effects of our applied theater-based gender-sensitization training on police officers. The unit of intervention for this study is a police station – denoted s – or an officer – denoted i. We randomly allocated stations into a treatment or control condition. Of the 419 stations in the

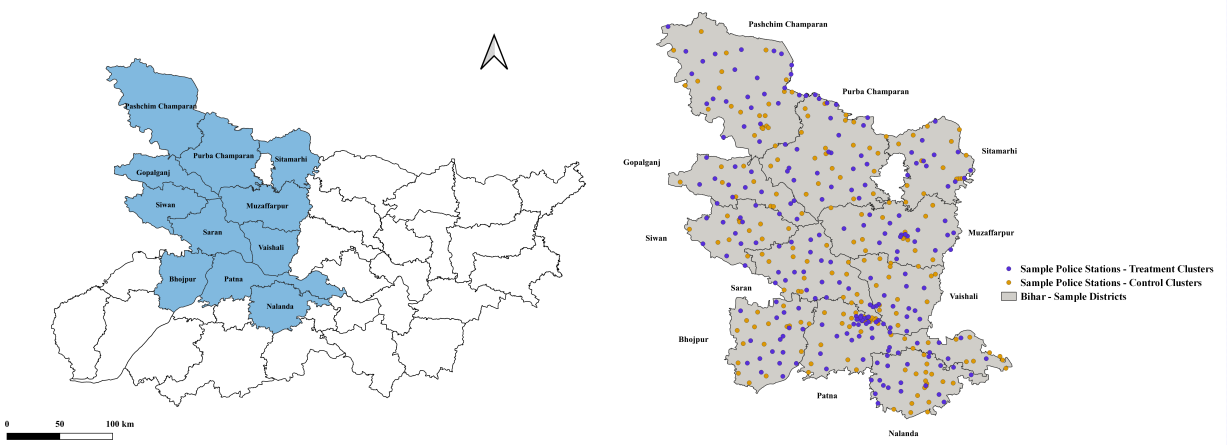
[14]Outposts are designed to serve as an extension of the main police station, and the personnel stationed there are responsible for maintaining law and order in the area. Outposts are typically located in areas that are far away from the main police station and are staffed by a few police personnel.

[15]https://www.nationalheraldindia.com/national/200-police-outposts-in-bihar-to-be-upgraded-shortly

study sample, in 217 stations selected officers were invited to attend the training program, while officers working in stations randomly assigned to control stations did not receive the training – 202 stations.

Within stations, all male officers at the rank of ASI, SI, Inspector, and SHO are invited to attend via a letter and through phone calls. The randomization was conducted for each district following the timeline of the completion of the baseline. Randomization was stratified by district and the total number of SHOs, Inspectors, SIs, ASIs, and PSIs posted at a station at the time of the baseline. In total, nearly 3,526 male police officers of these ranks completed the baseline survey. The stations in our sample and their treatment conditions are mapped in Figure 1. We approached the SPs of each district and obtained permission letters to conduct the survey and intervention activities in the district.

Figure 1: Study sample and distribution of police stations.



Note: The left panel highlights the study districts involved in the study. The right panel plots the police stations in the study and its treatment assignment.

4.2 Empirical Specification

To estimate the causal effect of the intervention at the individual-level, we estimate make use of ANCOVA implementing the following specification:

$$Y_{is} = \beta_0 + \beta_1 T_s + \gamma_r + Y_i^0 s + \epsilon_{ipd} \tag{1}$$

where Y_{is} is an index of overall GBV skills measured for officer i placed in station s during baseline. We also document the effects by attitudes toward GBV index GBV and technical skills, the sub-components of the overall GBV skills index. $Y_i^0 s$ is the baseline index for the respective post-intervention outcome of interest. In our estimations we include stratification fixed effects - including district given the staggered implementation across districts -, γ_r, in all specifications and cluster the standard errors at the station level. The components of the main skills targeted in the intervention are analyzed separately to better understand mechanisms. We present results for all the components of the attitudes toward GBV index which include victim-blaming, empathy, views of GBV, externalizing police responsibilities, and truthfulness of complaints. For technical skills we estimate the effect of the intervention on the ability to identify GBV, and legal knowledge.

Tn robustness tests, we add to our main specification police officer baseline controls such as age, caste and education. We also include a measure of police station conditions - police station facilities index as defined above, remoteness of a station and presence of conflict nearby a station. We also control for including enumerator fixed effects, and separately facilitator fixed effects.

To test whether changes in attitudes toward gender-based violence translate into improvements in workplace culture among female officers, we measure: (ii) work environment-distribution of work at the police station, (iii) job satisfaction, (iv) workplace well-being; and, (v) incidence of harassment.

4.3 Heterogeneity

We pre-specify heterogeneity analyses by baseline levels of the primary outcomes. We also examine heterogeneous treatment effects by strata defined by the gender–rank composition of officers and by the rural or urban status of police stations. In addition, we estimate heterogeneity by the presence of female senior officers and by transfer status. We exploit the fact that transfers are exogenous to treatment assignment (See Table A2) to shed light on underlying mechanisms.

First, we assess whether treated officers who transfer to untreated stations carry individual-

level changes with them. This allows us to distinguish whether observed effects reflect shifts in individual attitudes or changes in perceptions about what behaviors are accepted by senior peers. Second, by examining control officers who transfer to treated stations, we test for spillover effects and the diffusion of knowledge from treated high-ranking officers. Third, we evaluate the importance of leadership composition by comparing outcomes when untreated officers are exposed to stations with a single treated senior officer versus stations where all senior officers are treated. This analysis helps identify the role of reinforcement by peers and the extent to which full treatment saturation among senior ranks is necessary to generate change.

4.4 Validity of the Empirical Specification

Following randomization, we conduct balance tests using both officer-level and police-station–level baseline characteristics. Table 1 reports that treatment and control groups are statistically similar across the vast majority of observable characteristics, with only 4 out of 25 variables (16%) showing differences at conventional significance levels—consistent with random assignment.

Between baseline and endline, we successfully re-survey 73% of the original sample. A total of 955 officers attrit from the study sample—487 from the control group and 468 from the treatment group. The primary reason for attrition is administrative transfers of officers to districts outside the study area, a common feature of policing in this context. A small number of officers left the sample due to suspension, retirement, or ill health. We show that across treatment and control stations, there is a small and economically meaningless difference in attrition but, the characteristics of officers being transferred do not differ across treatment arms. As an additional robustness check, we compute Lee bounds to assess the sensitivity of our main estimates to differential attrition (Kremer et al., 2009; Baird et al., 2011; Drexler et al., 2014; Fiala et al., 2022).

Table 1: Balance at Baseline

	Control Mean	Treatment Mean	Difference (2)−(1)	SE	P-value
Panel A: Demographics					
Age	47.450	47.414	-0.010	0.315	0.975
Years of service (force)	22.575	22.192	-0.356	0.393	0.365
Years of service (station)	1.234	1.168	-0.059	0.043	0.173
Married	0.906	0.897	-0.007	0.010	0.467
Panel B: Caste					
Caste: SC	0.184	0.186	0.003	0.013	0.819
Caste: ST	0.029	0.034	0.004	0.006	0.471
Caste: OBC	0.393	0.367	-0.027	0.016	0.102
Caste: General	0.347	0.372	0.025	0.016	0.130
Panel C: Education					
Education: 10th	0.337	0.318	-0.020	0.016	0.201
Education: Plus 2	0.165	0.177	0.013	0.013	0.321
Education: Diploma	0.011	0.009	-0.002	0.003	0.636
Education: Some college	0.001	0.003	0.002	0.002	0.315
Education: B.A.	0.404	0.429	0.025	0.017	0.126
Education: Post Graduate	0.081	0.064	-0.016*	0.009	0.066
Panel D: Rank					
Rank: ASI	0.458	0.447	-0.013	0.017	0.452
Rank: SI	0.379	0.390	0.015	0.016	0.368
Rank: PSI	0.061	0.062	-0.000	0.008	0.965
Rank: Inspector	0.003	0.005	0.002	0.002	0.400
Rank: SHO	0.098	0.096	-0.003	0.010	0.729
Panel E: Police Station Characteristics					
PSFS General Index	0.158	0.178	0.037	0.028	0.186
PSFS M-F Segregation Index	0.044	-0.057	-0.096***	0.027	0.000
Panel F: Baseline Outcome Indices					
GBV Skills Index	0.000	0.058	0.060*	0.033	0.072
Attitudes Index	0.000	0.046	0.046	0.033	0.171
Technical Skills Index	0.000	-0.016	-0.012	0.033	0.719
Number of officers	1676	1850	3526		

Notes: Standard errors clustered at the station level. Each row reports results from a separate OLS regression of the baseline characteristic on the treatment indicator, absorbing strata fixed effects. The difference column reports the coefficient on treatment with stars denoting significance of the p-value. Sample restricted to officers who completed both baseline and endline surveys. Source: Baseline survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5 Results

5.1 First Stage: Empathy and Perspective Taking Improved

Using a short reflection survey administered on the final day of the training, we examine whether the experiential learning component increased empathy and perspective-taking—two central channels through which the Theatre of the Oppressed approach seeks to reshape attitudes and norms.

We construct a standardized index from validated items capturing empathetic concern (e.g., feeling protective when someone is taken advantage of, feeling touched by others' misfortune, describing oneself as soft-hearted) and perspective-taking (e.g., looking at both sides of a disagreement, imagining another person's viewpoint before criticizing them). The index also incorporates reverse-coded items reflecting emotional detachment or resistance to considering others' perspectives.

Table 2 shows that officers in treated stations exhibit statistically significant increases in this index. These effects reflect a greater willingness to understand others' emotional states and stronger affective responses to unfair treatment.

Several features of these findings are noteworthy. First, the measures capture general psychological constructs rather than attitudes specific to gender-based violence. This suggests that the intervention did not merely activate compliance with GBV-related protocols, but instead shifted broader socio-emotional parameters that govern interpersonal interactions. Overall, the evidence indicates that the training altered other-regarding preferences and reflective decision-making in ways that could shape discretionary behavior within police stations. In the next section, we examine whether these shifts in empathy and perspective-taking translate into improvements in GBV-related skills.

Table 2: The Role of Training on Reflection Index

	Reflection Index	
Treatment	0.090**	0.085**
	(0.043)	(0.043)
Control mean	0.000	0.000
Strata FE	Yes	Yes
Station controls	No	Yes
Number of Clusters	413	413
Number of Officers	2830	2830

Notes: Standard errors are clustered at the police station level in parentheses. The dependent variable is a standardized reflection index constructed from 14 items (5 items reverse-coded). All specifications include strata fixed effects. Column (2) adds station controls (general infrastructure, female infrastructure, and male-female segregation indices). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Male Officers' Gender-Based Violence Attitudes and Technical Skills

Table 3 presents the main treatment effects on endline outcomes.The expressive arts training significantly improves officers' GBV skills, attitudes, and technical skills by about 0.1 standard deviations relative to the control group. All specifications control for the baseline value of the outcome and include strata, district, and officer-strength fixed effects, with standard errors clustered at the station level. The sample consists of officers who completed both baseline and endline surveys, including those who transferred across stations. Overall, the results indicate positive and statistically significant impacts of training across all outcome domains. These results are robust to social desirability, controls, enumerator fixed effects (See Table **??**).

Table 3: Treatment effects on endline outcomes with multiple hypothesis testing corrections

	GBV Skills Index	Attitudes Index	Technical Skills Index
Treatment	0.104**	0.104**	0.086**
	(0.041)	(0.042)	(0.043)
Control mean	-0.033	-0.031	-0.024
Strata FE	Yes	Yes	Yes
Baseline control	Yes	Yes	Yes
Number of clusters	417	417	417
Number of officers	2571	2571	2571
Adjusted p-values			
Unadjusted p-value		0.014	0.045
Romano-Wolf		0.002	0.004
Bonferroni		0.028	0.090

Notes: Standard errors clustered at the station level (in parentheses). Regressions include only officers who completed both baseline and endline surveys, including those who switched between control and treatment stations. All regressions include district and officer strength fixed effects and the baseline value of the outcome. All indices created using inverse covariance weighting (swindex) and normalized to the control-group mean. The GBV Skills Index is excluded from the correction family as it is a composite of the Attitudes and Technical Skills indices. Romano-Wolf p-values control the familywise error rate using 1,000 bootstrap replications (seed 12345). Bonferroni p-values multiply the unadjusted p-value by $K = 2$. Source: Baseline and endline surveys; police station facility survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Mechanisms. Table 4 reports heterogeneous treatment effects by officers' transfer status between baseline and endline and provides insight into the mechanisms through which the intervention operates.

Relative to the omitted group—control officers who remained in control stations—treated officers that did not move and started and treated stations exhibit positive and statistically

Table 4: Differential Treatment Effects by Transfer Status

	GBV Skills Index	Attitudes Index	Technical Skills Index
Treatment stayer	0.132***	0.120**	0.125***
	(0.046)	(0.047)	(0.047)
T $\rightarrow$ C mover	0.159**	0.159**	0.062
	(0.070)	(0.072)	(0.087)
C $\rightarrow$ T mover	0.220***	0.144*	0.212***
	(0.078)	(0.084)	(0.077)
p-value: T Stayer vs Mover	0.691	0.582	0.463
p-value: C Stayer vs Mover	0.005	0.086	0.006
Strata FE	Yes	Yes	Yes
Baseline outcome	Yes	Yes	Yes
Number of clusters	417	417	417
Number of officers	2571	2571	2571

Standard errors clustered at the station level (in parentheses). Regressions include only officers who completed both baseline and endline surveys, and examine heterogeneity by transfer status. The omitted category is officers in control stations at baseline who did not move between baseline and endline to a treated station. Treatment stayers are officers who remained in treatment stations throughout. T $\rightarrow$ C movers are officers who were initially assigned to treatment stations but transferred to control stations between baseline and endline. C $\rightarrow$ T movers are officers who were initially assigned to control stations but transferred to treatment stations between baseline and endline. The bottom row reports the p-value testing whether these effects differ significantly. All regressions include district and officer strength (at the station level) fixed effects. All regressions include the baseline value of the outcome. All indices created using inverse covariance weighting (swindex) and normalized to the control-group mean. Source: Baseline and endline surveys; police station facility survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

significant gains across all outcomes: 0.132 ($p < 0.01$) for the GBV Skills Index, 0.120 ($p < 0.05$) for the Attitudes Index, and 0.125 ($p < 0.01$) for the Technical Skills Index. Officers who moved from treated to control stations also display improvements in GBV skills (0.159, $p < 0.05$) and attitudes (0.159, $p < 0.05$), indicating that gains persist even when officers are no longer exposed to treated peers. This pattern is consistent with a preference-based channel, as it suggests that training altered officers' internal attitudes and competencies rather than merely their behavior in response to local peer environments.

In contrast, officers who moved from control to treated stations experience sizable improvements, particularly in GBV skills (0.220, $p < 0.01$) and technical skills (0.212, $p < 0.01$), with weaker evidence for attitudes (0.144, $p < 0.10$). These effects point to a social-norm or peer-exposure mechanism: officers who were not directly treated nonetheless improve when they enter treated environments, consistent with learning, imitation, or shifts in perceived expectations generated by trained colleagues. Supporting this interpretation, tests comparing treated stayers and treated movers show no statistically significant differences, suggesting that treatment effects travel with individuals, while tests among control officers reveal significant differences, indicating that exposure to treated peers generates gains. Moreover, the larger effect sizes among officers who move to treated stations suggest that peer effects and station-level norms play a key role in amplifying treatment effects.

These patterns help distinguish mechanisms: persistence of effects after leaving treated stations is consistent with preference change, whereas improvements among control officers entering treated stations are consistent with norm diffusion and peer reinforcement. Overall, the transfer-based heterogeneity analysis suggests that the intervention operates through a combination of internalized preference change, social-norm formation, and technical skill acquisition, with transfers allowing us to separately identify individual persistence versus environment-driven effects.

To further shed light on these mechanisms, we examine whether treatment effects vary by baseline individual attitudes, the baseline attitudes of peers within training groups, officer rank, and the share of female officers in the station. These heterogeneity analyses

allow us to assess whether impacts are stronger among officers with initially more regressive views, whether peer composition shapes treatment effectiveness, and whether hierarchical position or gender representation conditions the transmission of training effects.

Table 5 examines whether the effectiveness of training depends on the baseline skill level of officers' training peers. The treatment coefficient captures the effect for officers whose training peers had baseline skills that were worse than or equal to their own, while the interaction term indicates how the treatment effect differs when peers had higher baseline skills. For officers whose peers were worse than or equal to them, training produces positive and statistically significant gains across all outcomes: 0.15 ($p < 0.01$) for the GBV Skills Index, 0.16 ($p < 0.01$) for the Attitudes Index, and 0.12 ($p < 0.05$) for the Technical Skills Index.

By contrast, the interaction terms are negative for GBV skills and attitudes, indicating weaker effects when officers trained alongside higher-skilled peers; this difference is marginally significant for GBV skills (interaction = -$0.26, p$¡0.10), while the interaction for attitudes (-0.18) is not statistically significant. For technical skills, the interaction is small and positive (0.01), and the implied treatment effects are similar across peer quality groups. A direct test of equality across peer quality groups is marginally significant for GBV skills (p = 0.093), but fails to reject the null for attitudes (p = 0.259) and technical skills (p = 0.953).

This pattern suggests that training is most effective for officers who are not dominated by higher-performing peers during training. This is consistent with mechanisms such as participation intensity, confidence, or learning-by-doing: when peers are stronger, weaker officers may defer participation or rely on others, reducing skill acquisition. In contrast, when officers train with peers of similar or lower baseline ability, they may engage more actively, resulting in larger gains.

Table 5: Does training work differently when peers are better vs. worse than the officer?

	GBV Skills Index	Attitudes Index	Technical Skills Index
Treatment (peers worse/equal than officer)	0.15***	0.16***	0.12**
	(0.04)	(0.04)	(0.05)
Peers better than officer (GBV)	0.11	-0.01	-0.13
	(0.15)	(0.15)	(0.17)
Treatment × Peers better (GBV)	-0.26*	-0.18	0.01
	(0.16)	(0.16)	(0.18)
TE: peers worse/equal to officer	0.151	0.158	0.116
SE: peers worse/equal to officer	0.044	0.045	0.045
TE: peers better than officer	-0.114	-0.022	0.127
SE: peers better than officer	0.151	0.153	0.177
p-value: difference by peer quality	0.093	0.259	0.953
Control mean	-0.03	-0.03	-0.02
Strata FE	Yes	Yes	Yes
Baseline outcome	Yes	Yes	Yes
Clusters	417	417	417
Officers	2571	2571	2571

Notes: Standard errors clustered at the station level (in parentheses). Positive coefficients indicate improvements. For each officer with observed training group identifier, we calculate their peer baseline as the leave-one-out mean baseline index value of other officers in their training group (excluding the officer themselves). The indicator equals 1 if this peer average is higher than the officer's own baseline, 0 otherwise. Officers with no training group data are coded as 0. The treatment coefficient shows the treatment effect for officers whose training peers had baseline skills worse than or equal to the officer's own baseline (or who had no training group). The interaction term shows how much the treatment effect differs for officers whose training peers had higher baseline skills than them. The bottom rows report the full treatment effect for each group: officers whose peers were worse/equal vs. officers whose peers were better. The p-value tests whether these two treatment effects are statistically different from each other. All regressions include district and officer strength (at the station level) fixed effects. All regressions include the baseline value of the outcome. All indices created using inverse covariance weighting (swindex) and normalized to the control-group mean.
Source: Baseline and endline surveys; police station facility survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Finally, we find no evidence that treatment effects vary systematically by officer rank, station urbanicity, or the share of female constables. The absence of heterogeneity along these dimensions suggests that the observed impacts are unlikely to be driven by differential monitoring, supervision, or gender composition across stations, and instead reflect changes that operate broadly through changes in individual attitudes.

5.3 Workplace Environment for Female Officers

In Table 6, we examine whether training senior officers generated spillovers to the workplace experiences of junior female constables. Eight months after the intervention, female officers in treated stations report significantly improved workplace environments, with the overall workplace environment index declining by 0.39 standard deviations ($p < 0.05$), indicating more favorable conditions. Reports of harassment also decrease by 0.06 ($p < 0.10$), while anxiety measures show no statistically significant change. Importantly, stations with a higher share of trained officers exhibit additional improvements: greater training saturation is associated with lower anxiety (0.42, $p < 0.10$) and more positive perceptions of male colleagues (0.18, $p < 0.05$). These findings suggest that the intervention produced institutional spillovers that improved organizational climate for women, consistent with a mechanism in which training top officers shifts norms within stations rather than affecting only directly treated officers. In particular, the fact that outcomes for female constables improve despite not being direct recipients of the intervention points to a station-level channel operating through supervisory practices and peer interactions.

Table 6: Treatment effects on Female Constables

Main Index	Panel A			Panel B		
	Anxiety Score	Perceptions Towards Male Officers	Reported Harassment	Anxiety Score	Perceptions Towards Male Officers	Reported Harassment
Treatment	-0.39**	0.09	-0.06*			
	(0.15)	(0.06)	(0.04)			
Share Trained Officers				-0.42*	0.18**	-0.07
				(0.22)	(0.09)	(0.05)
Control mean	2.53	0.00	0.00	2.53	-0.04	-0.01
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	366	366	366	366	366	366
Number of officers	1810	1810	1803	1810	1810	1803

Notes: Standard errors clustered at the station level (in parentheses). All regressions include district and officer strength (at the station level) fixed effects. **All indices created using inverse covariance weighting, and normalised for control group mean.** Treatment is a dummy variable that equals 1 if the female police officer surveyed at endline works in a police station assigned to the treatment group and 0 otherwise. Share of Trained Officers represents the proportion of senior police officers in a given police station who received treatment. On average, 38.43% of senior officers across all police stations were trained. Source: Female constables' survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.4 Male Officers' Performance in Gender-Based Violence Cases

To assess whether training translated into changes in real-world policing behavior, we analyze outcomes from standardized decoy visits conducted approximately one year after the intervention in Table Table 7. Officers in treated stations are significantly less likely to dismiss complaints as a waste of time, claim cases are false, blame the victim, or ask unnecessary questions, and are more likely to refer complainants to appropriate services and explain available alternatives clearly. Because these measures capture observed officer conduct rather than self-reported attitudes, they provide strong evidence that the intervention improved actual case-handling behavior. The persistence of these effects one year after training further indicates that impacts are durable and reflect meaningful changes in professional practice rather than short-lived survey responses or experimenter demand effects. Taken together with the workplace spillover results, the decoy evidence suggests that the intervention operates through both internalized changes in officers' preferences and externally reinforced norms within stations, ultimately translating attitudinal shifts into observable behavioral improvements.

Table 7: Decoy Results

	Waste of time	Cases false	Blames victim	Referred to WPS	Explained alternative	Asks unnecessary questions
Treatment x GBV case	-0.07***	-0.08***	-0.15***	0.07***	0.08*	-0.14***
	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)
Control mean	0.06	0.15	0.72	0.09	0.21	0.63
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	183.00	183.00	183.00	183.00	183.00	183.00
Number of officers	549.00	549.00	549.00	549.00	549.00	549.00

Notes: Standard errors clustered at the station level (in parantheses).$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In the Appendix, to further shed light on the change of attitudes mechanism, we analyze whether effects are driven by officers with less progressive views at baseline.

6 Conclusion

Public institutions frequently fail the most vulnerable not only because of limited resources or weak incentives, but also because of the beliefs and biases of frontline state agents. In the context of policing, discretion is central: officers determine whether complaints are registered, how victims are treated, and whether the law is applied consistently. When these decisions are shaped by entrenched gender norms, access to justice becomes unequal.

This paper studies whether reshaping the attitudes and skills of senior police officers through an immersive expressive arts training can improve institutional performance. We evaluate a large-scale clustered randomized controlled trial covering 419 police stations and approximately 2,000 senior officers in Bihar, India. The intervention draws on experiential learning and perspective-taking exercises designed to surface and challenge implicit biases within a highly hierarchical organization.

We document three main findings. First, the training increases empathy and perspective-taking among treated officers, shifting general psychological constructs rather than merely activating GBV-specific compliance. Second, these shifts translate into improvements in attitudes towards GBV, technical knowledge and workplace culture, including reductions in harassment reported by female constables. Moreover, we quantify significant spillover effects to non treated officers highlighting the importance of changing the norm of officers at the top. Third, and most importantly, we observe meaningful improvements in citizen-facing behavior. Using decoy audits conducted one year after the intervention, we find that officers in treated stations are less likely to dismiss complaints, blame victims, or engage in inappropriate questioning, and more likely to provide clear procedural guidance and referrals. Effects are strongest among officers with more regressive baseline attitudes, consistent with the intervention relaxing a bias-based constraint on discretionary enforcement.

These findings contribute to the literature on state capacity by demonstrating that bias within state agents is not only consequential but malleable. Unlike reforms that rely on changing personnel composition or creating parallel institutions, our intervention operates within the existing hierarchy, targeting decision-makers who shape both case outcomes and workplace norms. The results suggest that socio-emotional skills, attitudes, norms and institutional culture are integral components of public sector performance.

More broadly, the evidence highlights that improving governance may require addressing not only incentives and information, but also the social and psychological foundations of bureaucratic decision-making. In settings where private alternatives to public safety are limited, enhancing the quality and fairness of state institutions is particularly consequential. Interventions that shift perspective-taking and other-regarding preferences among senior officials may therefore represent a scalable pathway toward more equitable access to justice.

References

Adams-Prassl, A., Huttunen, K., Nix, E., and Zhang, N. (2024). Violence against women at work. *The Quarterly Journal of Economics*, 139(2):937–991.

Alan, S., Baysan, C., Gumren, M., and Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, 136(4):2147–2194.

Amaral, S., Borker, G., Fiala, N., Kumar, A., Prakash, N., and Sviatschi, M. M. (2025). Sexual harassment in public spaces and police patrols: Experimental evidence from urban india. *The Quarterly Journal of Economics*, 140(4):3191–3231.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Andries, M., Bursztyn, L., Chaney, T., Djourelova, M., and Imas, A. (2024). In their shoes: Empathy through information. Technical Report Working Paper 32569, National Bureau of Economic Research.

Ashraf, N., Bryan, G., Delfino, A., Holmes, E., Iacovone, L., Meyer, C., and Pope, A. (2024). Learning to see a world of opportunities. *Unpublished Manuscript. Available at: https://navaashraf. com/wp-content/uploads/2024/09/learning_to_see_a_world_of_opportunities-1. pdf*.

Baird, S., McIntosh, C., and Özler, B. (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly journal of economics*, 126(4):1709–1753.

Banerjee, A. V., Chattopadhyay, R., Duflo, E., Keniston, D., and Singh, N. (2021). Improving police performance in rajasthan, india: experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy*, 13(1):36–66.

Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American economic review*, 99(4):1218–1244.

Boal, A. (2014). Theatre of the oppressed. In *The Improvisation Studies Reader*, pages 97–104. Routledge.

Boudreau, L. E., Chassang, S., Gonzalez-Torres, A., Heath, R., and of Economic Research, N. B. (2023). Monitoring harassment in organizations. Technical report, National Bureau of Economic Research.

Brar, R., Buehren, N., Papineni, S., and Sulaiman, M. (2025). Resisting social conformity pressure: Impact of a gender norms intervention for adolescents in somalia. *Journal of Development Economics*, page 103713.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American economic review*, 110(10):2997–3029.

Bursztyn, L., Haaland, I. K., Röver, N., and Roth, C. (2025). The social desirability atlas. Technical report, National Bureau of Economic Research.

Bursztyn, L. and Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9:131–153.

Canales, R., Santini, J. F., González Magaña, M., and Cherem, A. (2025). Shaping police officer mindsets and behaviors: Experimental evidence of procedural justice training. *Management Science*, 71(11):8995–9013.

Crowne, D. P. and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, 24(4):349–354.

Cunha, F., Hu, Q., Xia, Y., and Zhao, N. (2023). Reducing bullying: Evidence from a parental involvement program on empathy education. Technical report, National Bureau of Economic Research.

Dahl, G. B. and Knepper, M. M. (2021). Why is workplace sexual harassment underreported? the value of outside options amid the threat of retaliation. Technical report, National Bureau of Economic Research.

Das, J., Holla, A., Mohpal, A., and Muralidharan, K. (2016). Quality and accountability in health care delivery: audit-study evidence from primary care in india. *American Economic Review*, 106(12):3765–3799.

Datta, A. and Satija, S. (2020). Women, development, caste, and violence in rural bihar, india. *Asian Journal of Women's Studies*, 26(2):223–244.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American economic review*, 112(3):899–927.

Di Maio, M. and Fiala, N. (2020). Be wary of those who ask: a randomized experiment on the size and determinants of the enumerator effect. *The World Bank Economic Review*, 34(3):654–669.

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2):1–31.

Dube, O., MacArthur, S. J., and Shah, A. K. (2024). A cognitive view of policing*. *The Quarterly Journal of Economics*, 140(1):745–791.

Dube, O., MacArthur, S. J., and Shah, A. K. (2025). A cognitive view of policing. *The Quarterly Journal of Economics*, 140(1):745–791.

Fiala, N., Garcia-Hernandez, A., Narula, K., and Prakash, N. (2022). Wheels of change: Transforming girls' lives with bicycles.

Finan, F., Olken, B. A., and Pande, R. (2017). The personnel economics of the developing state. *Handbook of economic field experiments*, 2:467–514.

Folke, O. and Rickne, J. (2022). Sexual harassment and gender inequality in the labor market. *The Quarterly Journal of Economics*, 137(4):2163–2212.

Freire, P. (1970). *Pedagogy of the Oppressed*. Herder and Herder, New York. Translated by Myra Bergman Ramos.

Freire, P. (2020). Pedagogy of the oppressed. In *Toward a sociology of education*, pages 374–386. Routledge.

García-Hombrados, J., Martínez-Matute, M., and Villa, C. (2024). Specialised courts and the reporting of intimate partner violence: Evidence from spain. *Journal of Public Economics*, 239:105243.

Harris, D., Borcan, O., Serra, D., Telli, H., Schettini, B., and Dercon, S. (2022). Proud to belong: The impact of ethics training on police officers.

Harris, D., Borcan, O., Serra, D., Telli, H., Schettini, B., and Dercon, S. (2024). Proud to belong: The impact of ethics training on police officers in ghana.

Indian Institute of Population Sciences (2021). *National Family Health Survey (NFHS-5), India, 2019-21: Bihar*. MoHFW, Government of India, Mumbai: IIPS.

Indian Institute of Population Sciences (2022). *National Family Health Survey - 5: India Fact Sheet*. MoHFW, Government of India, Mumbai: IIPS.

Iyer, L. and Mani, A. (2012). Traveling agents: political change and bureaucratic turnover in india. *Review of Economics and Statistics*, 94(3):723–739.

Jolliffe, D. and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of adolescence*, 29(4):589–611.

Kremer, M., Miguel, E., and Thornton, R. (2009). Incentives to learn. *The Review of Economics and statistics*, 91(3):437–456.

Matavelli, I. (2024). We don't talk about boys: masculinity norms among adolescents in brazil. *University of British Columbia, mimeo*.

Miller, A. R. and Segal, C. (2019). Do female officers improve law enforcement quality? effects on crime reporting and domestic violence. *The review of economic studies*, 86(5):2220–2247.

National Crime Records Bureau, M. o. H. A. (2023). *Crime in India 2022: Volume I*. Government of India.

National Herald India (2022). 200 police outposts in bihar to be upgraded shortly. [Accessed on April 4, 2024].

Page, L. and Pande, R. (2018). Ending global poverty: Why money isn't enough. *Journal of Economic Perspectives*, 32(4):173–200.

Rehfisch, J. M. (1958). A scale for personality rigidity. *Journal of Consulting Psychology*, 22(1):11.

Sharma, K. (2022). Tackling sexual harassment: Experimental evidence from india. *Working Paper*.

Stewart, R., Wright, B., Smith, L., Roberts, S., and Russell, N. (2021). Gendered stereotypes and norms: A systematic review of interventions designed to shift attitudes and behaviour. *Heliyon*, 7(4).

Sukhtankar, S., Kruks-Wisner, G., and Mangla, A. (2022). Policing in patriarchy: An experimental evaluation of reforms to improve police responsiveness to women in india. *Science*, 377(6602):191–198.

Sviatschi, M. M. and Trako, I. (2024). Gender violence, enforcement, and human capital: Evidence from women's justice centers in peru. *Journal of Development Economics*, 168:103262.

Tata Trusts (2023). India justice report: Ranking states on police, judiciary, prisons and legal aid (2022). Technical report, Tata Trusts, New Delhi, India. In collaboration with Centre for Social Justice, Common Cause, Commonwealth Human Rights Initiative, DAKSH, TISS-Prayas, and Vidhi Centre for Legal Policy.

	Status		Treatment	Control	Total
In-study sample					
(1)	Endline and baseline		1382	1189	2571
(1A)	Not transferred		1046	872	1918
(1B)	Transferred		336	317	653
		i. From control to control	-	148	148
		ii. From treatment to treatment	161	-	161
		iii. From control to treatment	-	169	169
		iv. From treatment to control	175	-	175
Mid-way enrollment sample					
(4)	Baseline only		468	487	955
(5)	Endline only		363	277	640
(5A)	Endline and training		152	15	167
Total (1)+(4)+(5)			**2213**	**1953**	**4166**

Appendix A - Tables and Figures

Table A.1: Summary - police officers in sample districts

Officer Rank	Male	Female
Station Head Officer (SHO)	388	6
Inspector	42	0
Sub-Inspector of Police (SI)	1549	230
Assistant Sub-Inspector (ASI)	1627	39
Head Constable	400	84
Writer Constable	349	30
Constable	4250	1596
Total	**8605**	**1985**

Notes: This table summarizes police officers across various designations in the sample districts at the time of the baseline survey.

Table A2: Transfers are orthogonal to treatment assignment

	Moved across treatment arms		
Treatment	-0.018	-0.014	-0.016
	(0.014)	(0.014)	(0.014)
Strata FE	Yes	Yes	Yes
Officer controls	No	Yes	Yes
Station controls	No	No	Yes
Number of Clusters	417	417	417
Number of Officers	2571	2571	2571

Notes: Robust standard errors clustered at the police station level in parentheses. The dependent variable equals one if the officer moved between treatment and control stations between surveys. All specifications include district fixed effects and baseline officer strength controls. Column (2) adds officer controls (age, years of service, education, rank, caste, marital status). Column (3) adds station controls (infrastructure, facilities, officer ratios). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix B - Main Outcomes and Indices Construction

B.1 Procedure for index construction

The indices for our primary and secondary outcomes are constructed by aggregating the responses to several individual questions following the procedure of Anderson (2008), implemented via the `swindex` command in Stata. Specifically, for each index, the individual component variables are first standardized to have mean zero and unit standard deviation using the control group mean and standard deviation as the reference. The standardized components are then aggregated into a summary index using a weighted average, where the weights are derived from the inverse of the covariance matrix of the standardized components. This weighting scheme assigns relatively less weight to components that are highly correlated with one another, thereby reducing redundancy and increasing the efficiency of the index. The resulting index is itself re-standardized to have mean zero and unit standard deviation in the control group, so that estimated treatment effects can be interpreted in standard deviation units.

B.2 Primary outcomes

B.2.1 Victim blaming index

Victim blaming involves the devaluation of the circumstances of an incident occurring to an individual. It also involves the act of holding an individual responsible for the circumstances that led to such crime victimization. Blaming individuals as a result of the circumstances or behaviors associated with crime victimization may discourage victims from reporting and engaging with criminal investigations - for instance, by lowering self-esteem, confidence or even, by decreasing the capacity to recall the circumstances of the

incidence. As a result, victim-blaming ultimately impacts police performance by lowering the quality of questioning, investigation, probability of having an FIR filled. The baseline victim blaming index aggregates the following 7 survey questions:

Q202a. It is justified for a husband to hit his wife if she goes out without telling him.
Q202b. It is justified for a husband to hit his wife if she neglects the children.
Q202c. It is justified for a husband to hit his wife if she burns the food.
Q202d. It is justified for a husband to hit his wife if she argues with the husband.
Q202e. It is justified for a husband to hit his wife if she refuses to have sex with the husband.
Q405. (Referring to vignette) Out of 10 such cases, in how many cases the woman faced such a situation because she did not behave in a socially acceptable manner?
Q602b. (Referring to vignette) To what extent do you agree that this incident happened because Anita was traveling alone?

B.2.2 Technical skills index

This is the measure of an officer's understanding of what the correct procedures are to be employed in cases of GBV. These procedures include soft and legal/criminal knowledge. The baseline technical skills index aggregates the following 8 survey questions:

Q301. Can the relatives of a victim who has experienced domestic violence file a police complaint on her behalf?
Q302. (Referring to vignette) With whom does the burden of proof lie?
Q303. If a married woman is evicted from her 'matrimonial house' but has not been subjected to any other physical abuse, does that constitute domestic violence?
Q304. Should the police assist the victim in finding shelter if she does not feel safe at her place?
Q305. A woman accuses an unknown man of having verbally abused her in a public space. Is this a chargeable offense?
Q305a. If yes, under what section of the IPC will you book him?
Q306. Your colleague, another police officer, has voluntarily revealed the name and the whereabouts of a rape victim to a local media outlet. Is this a chargeable offence?
Q306a. If the answer to the above question is yes, will you book him? If yes, under which section of the IPC shall you book him?

B.2.3 Empathy index

The empathy index measures feelings of empathetic concern among the officers and is based on the Basic Empathy Scale (BES) developed by Jolliffe and Farrington (2006). The baseline empathy index aggregates the following 8 survey questions:

Q501. I tend to have very strong opinions about morality.
Q502. I find it easy to put myself in somebody else's shoes.
Q503. I am good at predicting how someone will feel.
Q504. I am able to make decisions without being influenced by others' feelings.
Q505. I can tune into how someone else feels rapidly and intuitively.
Q506. I can usually appreciate the other person's viewpoint, even if I don't agree with it.
Q601e. (Referring to vignette) I can imagine what it must be like to be in Radha's place.
Q602c. (Referring to vignette) I can imagine what it must be like to be in Anita's place.
Q801e. (Referring to vignette) I can imagine what it must be like to be in Preeti Manjhi's/Varsha Gupta's place.

B.2.4 Acceptance towards GBV index

This outcome measures views held by the officers about what forms of abuse constitute a GBV act, what should be tolerated by women/not perpetrated by women, and in what ways they should be helping themselves (or not). Less equal norms around what women should tolerate and act upon regarding GBV could lower police officers' efforts and performance. The baseline attitudes index aggregates the following 4 survey questions:

Q402. (Referring to vignette) Out of 10 such cases, in how many cases is recommending a compromise enough and there is no need to file an FIR?
Q408. Cases related to women receive too much attention by the police, relative to other crime and law and order issues.
Q601a. (Referring to vignette) Thinking about the story, do you think that this consists of abusive behavior by Bablu?
Q601d. The independence enjoyed by women these days is causing such problems.
Q801a. Now, thinking about the story, do you think that Vijay Kumar is at fault in any way?

B.3 Secondary outcomes

B.3.1 Externalising police responses index

This outcome measures the extent to which the police externalise their responsibilities onto colleagues and/or victims in order to remove themselves from the process of justice. The baseline index aggregates the following 6 survey questions:

Q201a. The woman does not report to the police because the police cannot help her and it is an internal matter between the husband and the wife.
Q201b. The woman does not report to the police because such incidents are very common and happen too often.
Q201c. The woman does not report to the police because she fears being beaten up by her husband again if she reports to the police.
Q601b. (Referring to vignette) Do you think this is a problem that the police can help

with?
Q602d. Do you think the police can register a case against this man based on the complaint?
Q801b. (Referring to vignette) Do you think this is a problem that the police can help with?

B.3.2 Truthfulness of complaints index

Truthfulness of complaints refers to relying on instincts and intuition to judge the veracity of a criminal complaint quickly. This is in contrast with relying on facts and logical rationale to conduct decisions. Such beliefs about the veracity of complaints are often referred to as 'gut feelings' or 'hunches' and are likely a form of decision-making that relies on experience with past cases. The use of information based on experience, while relevant in any profession, may also lead to an overestimation of false cases if such perceptions are formed based on biased beliefs about victims and types of cases. Such overestimation may lead to a quick dismissal of reports or lower effort in police investigations. The baseline index aggregates the following 5 survey questions:

Q401. (Referring to vignette) Out of 10 cases how many such complaints are false?
Q403. (Referring to vignette) In how many such cases out of 10 are women pressing false charges related to sexual harassment in order to gain unfair advantage in land dispute cases?
Q404. (Referring to vignette) Out of 10 cases how many such complaints are false?
Q406. (Referring to vignette) Out of 10 such cases, in how many cases is the woman trying to frame the man after a fall-out of the relationship?
Q407. The complaint is usually more believable if the woman is also accompanied by her relatives rather than when she comes alone.
Q601c. (Referring to vignette) The complaint filed by Radha is true.
Q602a. (Referring to vignette) To what extent do you believe that Anita is speaking the truth?
Q801c. (Referring to vignette) The complaint filed by Reena is true.
Q801d. (Referring to vignette) This is a typical case that I have seen being reported to the Police Station where a woman tries to frame a man.

Appendix F - Research Ethics

The research had 7 study branches that required interactions with human subjects: baseline, endline and reflection survey of police officers in thanas, police station facilities surveys, female constable surveys and decoy surveys. All activities were developed by the researchers and implemented by DAI Research and Advisory Services.

Informed consent, safety protocol and withdrawal systems for enumerators. Prior to the initiation of the study, DAI conducted a training exercise with all enumerators. During the training, all individuals went through 7 days of training in which the research team in-

formed potential enumerators of the study and its aims and objectives. Enumerators who initially agreed to participate in the surveys were welcome to retire anytime they wish to without any financial consequences or any form of penalties. This aspect was reiterated to the enumerators not only during the training sessions but also on multiple counts during the duration of field work. The enumerators were also trained and provided with a field protocol document. This guidebook was written to be consulted if and when needed during fieldwork, and served as a guideline for particular training sessions. It detailed day to day activities during data collection, especially pertaining to enumerator observation survey, provided an overview of safety risks, outlined mitigation strategies and highlighted key support resources available to field staff during and in some cases after completion of fieldwork activities. Enumerators during the field work activities will also have at their disposal a field supervisor at their disposal to discuss openly any issue they may bear. Enumerators also engaged in group sessions to generate a dialogue of any potential unpleasant experiences they may have encountered. This was done in order to facilitate a better way to manage any potential unpleasant aspects of conducting the decoy survey. Enumerators were not aware of the randomization and the treatment assignment at each police station. Enumerators were also in permanent contact - via mobile phone - with the project assistants in order to discuss any potential harmful situation. This allowed for effective monitoring and ensuring the safety of enumerators. For the questionnaires for women (female constables and wives surveys) - we had set a specific ethics protocol as the survey instrument contained questions regarding violence. This could have possibly triggered trauma for respondents who had witnessed such violence first hand. During the survey training sessions, all associated risks along with resources they possess for help and risk mitigation will be thoroughly explained and discussed in plenary sessions and enumerators will be encouraged to discuss any concerns with the lead investigators of this project. With the help of DAI and trained theatre professionals, we developed exercise training modules to simulate situations that could arise at the police stations while enacting the cases for the decoy exercise. We also received prior permissions from the SPs of the various districts.

Online Appendix - Debiasing Law Enforcement Officers: Evidence from an Expressive Arts Intervention in India

Sofia Amaral* Girija Borker† Nishith Prakash‡ Maria Micaela Sviatschi§

March 26, 2026

Contents

1 Introduction **2**

2 Ethics and IRB Oversight **2**

3 Sample Selection **3**
3.1 Male officers . . . 3
3.2 Female Constables . . . 6
3.3 Audit Exercise . . . 6

4 The Intervention **6**
4.1 Ideation . . . 7
4.2 Designing the Training Curriculum . . . 7
4.3 Designing the Training Workshop . . . 8
4.4 Developing a team of facilitators . . . 11
4.5 Implementation of the intervention . . . 12
4.6 Monitoring of the intervention . . . 13

5 The Decoy Experiment **14**
5.1 Objective . . . 14
5.2 Sampling . . . 14
5.3 Standardized Victims (SVs) . . . 14
5.4 Implementation Protocols . . . 15
5.5 Mitigating the Potential Risks . . . 15

*South Asia Gender Innovation Lab, World Bank. Email: s.amaral@worldbank.org
†DIME, World Bank. Email: gborker@worldbank.org
‡Northeastern University, CESifo, IZA, CReAM, HiCN, & GLO. Email: n.prakash@northeastern.edu
§Princeton University. Email: msviatschi@princeton.edu

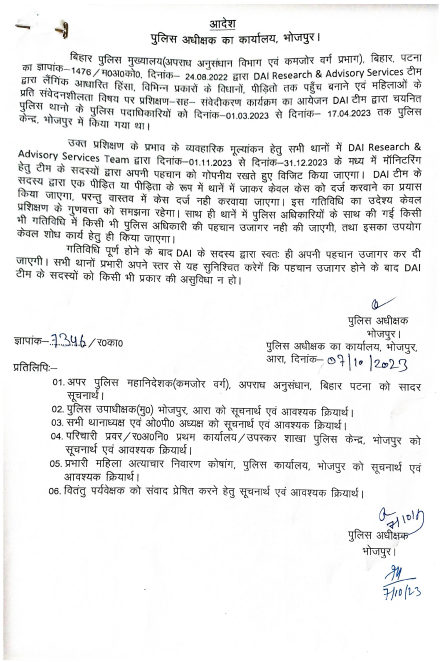
1 Introduction

This appendix details the various aspects of the study. The layout of the appendix goes thus: first, we elaborate on the details of the Institutional Review Board (IRB) approvals and protocols we followed during the study. Second, we go into the sample selection and attrition faced during the course of the study. Then we explain the working behind the curriculum that was used for training. Next, we briefly describe the decoy exercise we conducted at police stations where standardized victims (SVs) were asked to observe the functioning of the police in a business-as-usual setting. Afterwards, we briefly summarize the power calculations for the study. Next, we describe the various outcomes we measure during the analysis. Finally, we show the spatial figures we created for this study.

2 Ethics and IRB Oversight

This appendix describes the ethical considerations for this study. This study received approval from multiple IRBs in India and the United States. The Institutional Review Board of Princeton University provided primary oversight for the entire study. We also received IRB approval for the full study, including the interventions and data collection, from DAI Research and Advisory Services Pvt Ltd, which led the data collection. The study was also registered at the AEA Registry under the ID AEARCTR-0013270.

The IRB protocol for this study delineated practices to ensure the safety and protection of study participants. Subjects gave informed consent before participating in the baseline survey to identify officers who were eligible for the study. While consent from the officers was not sought to invite them to participate in the training, permission letters were obtained from the Superintendents of Police (SPs) in each district, which mandated the officers in the police stations under their respective jurisdictions to attend the training (an example is provided below).



आदेश

पुलिस अधीक्षक का कार्यालय, भोजपुर।

बिहार पुलिस मुख्यालय(अपराध अनुसंधान विभाग एवं कमजोर वर्ग प्रभाग), बिहार, पटना का ज्ञापांक–1476/म0अ0को0, दिनांक– 24.08.2022 द्वारा DAI Research & Advisory Services टीम द्वारा लैंगिक आधारित हिंसा, विभिन्न प्रकारों के विधानों, पीड़ितो तक पहुँच बनाने एवं महिलाओं के प्रति संवेदनशीलता विषय पर प्रशिक्षण–सह– संवेदीकरण कार्यक्रम का आयोजन DAI टीम द्वारा चयनित पुलिस थानो के पुलिस पदाधिकारियों को दिनांक–01.03.2023 से दिनांक– 17.04.2023 तक पुलिस केन्द्र, भोजपुर में किया गया था।

उक्त प्रशिक्षण के प्रभाव के व्यवहारिक मूल्यांकन हेतु सभी थानों में DAI Research & Advisory Services Team द्वारा दिनांक–01.11.2023 से दिनांक–31.12.2023 के मध्य में मॉनिटरिंग हेतु टीम के सदस्यों द्वारा अपनी पहचान को गोपनीय रखते हुए विजिट किया जाएगा। DAI टीम के सदस्य द्वारा एक पीड़िता या पीड़िता के रूप में थानें में जाकर केवल केस को दर्ज करवाने का प्रयास किया जाएगा, परन्तु वास्तव में केस दर्ज नही करवाया जाएगा। इस गतिविधि का उद्देश्य केवल प्रशिक्षण के गुणवत्ता को समझना रहेगा। साथ ही थानें में पुलिस अधिकारियों के साथ की गई किसी भी गतिविधि में किसी भी पुलिस अधिकारी की पहचान उजागर नही की जाएगी, तथा इसका उपयोग केवल शोध कार्य हेतु ही किया जाएगा।

गतिविधि पूर्ण होने के बाद DAI के सदस्य द्वारा स्वतः ही अपनी पहचान उजागर कर दी जाएगी। सभी थानों प्रभारी अपने स्तर से यह सुनिश्चित करेगें कि पहचान उजागर होने के बाद DAI टीम के सदस्यों को किसी भी प्रकार की असुविधा न हो।

पुलिस अधीक्षक
भोजपुर।

ज्ञापांक–7346/र0का0
पुलिस अधीक्षक का कार्यालय, भोजपुर,
आरा, दिनांक– 07/10/2023

प्रतिलिपि:–

01. अपर पुलिस महानिदेशक(कमजोर वर्ग), अपराध अनुसंधान, बिहार पटना को सादर सूचनार्थ।
02. पुलिस उपाधीक्षक(मु0) भोजपुर, आरा को सूचनार्थ एवं आवश्यक क्रियार्थ।
03. सभी थानाध्यक्ष एवं ओ0पी0 अध्यक्ष को सूचनार्थ एवं आवश्यक क्रियार्थ।
04. परिचारी प्रवर/र0अ0नि0 प्रथम कार्यालय/उपस्कर शाखा पुलिस केन्द्र, भोजपुर को सूचनार्थ एवं आवश्यक क्रियार्थ।
05. प्रभारी महिला अत्याचार निवारण कोषांग, पुलिस कार्यालय, भोजपुर को सूचनार्थ एवं आवश्यक क्रियार्थ।
06. वित्तु पर्यवेक्षक को संवाद प्रेषित करने हेतु सूचनार्थ एवं आवश्यक क्रियार्थ।

पुलिस अधीक्षक
भोजपुर।

7/10/23

Scanned with AnyScanner

Figure 1: *An example of a permission letter we obtained from each district SP for the intervention and data collection activities*

For the data collection activities involving female constables and spouses of officers in the sample, the survey team was composed exclusively of female enumerators. They were also trained in sensitivity to specific modules of sexual harassment, and were asked to elaborate on this during the process of obtaining informed

consent. All participation in the intervention and the data collection activities were completely voluntary, and participants could choose to withdraw their consent at any point of time during data collection.

Detailed protocols were followed to ensure the safety of the SVs during the decoy exercise. To mitigate potential risks in the decoy experiment, several safeguards were put in place. The safety of the SVs were ensured by informing the Station House officers (SHOs) of the relevant police stations about a potential decoy visit within a specified time frame, without revealing the identity or exact timing. The research team coordinated closely with the district SP to facilitate smooth implementation, and each visit was supervised by a dedicated field team. Confidentiality was strictly maintained, with the identities of both SVs and officers anonymized and all data securely stored on encrypted servers. To prevent the risk of fake case registrations, SVs were trained to clearly communicate the nature of the visit to officers and ensure that no official records (online or offline) were created. Additionally, psychological support and counselling services were made available to SVs before, during, and after their interactions to address any emotional or mental health concerns.

All research field staff signed confidentiality agreements and were trained in Princeton University's privacy protocols. Data was collected digitally using password-protected Android devices, with no use of paper forms. Survey data was encrypted upon completion and securely transferred to a password-protected server. To further protect participant privacy, all data was de-identified, with identifying information stored separately in a secure, restricted-access database. A secure, cloud-based storage system compliant with international data protection standards was used, with access limited to authorized personnel through strict user permissions and access controls.

3 Sample Selection

3.1 Male officers

Bihar has 1,604 stations with a total of 6,912 female officers and 78,203 male officers distributed across districts (BPRD, 2017). Within Bihar, we selected 12 study districts, namely, Patna, Bagaha, Bettiah, Motihari, Siwan, Saran, Gopalganj, Muzaffarpur, Vaishali, Sitamarhi, Nalanda, and Bhojpur, which have 419 police stations and cater to approximately 40 million households. This area was selected based on our consultations with senior Bihar Police officials because it has adequate police stations and number of police officers suitable to power our study. Our intervention was clustered at the station-level and we randomly allocated stations to be treatment or control stations. Selected officers from treated police stations attended the training program. officers working in stations randomly assigned to control stations did not receive the training. Within stations, all male officers at the rank of Assistant Sub-Inspector (ASI), Sub-Inspector (SI), and Police Sub-Inspector in Training (PSI) were selected to attend training sessions. We see from the table below that on average, the sample police station in the study had around eight police officers at the ASI, SI, and PSI ranks.

Table 1: Distribution of police stations and officers (baseline) by district

District	Police stations	Police officers (baseline)	Officers per PS
Bagaha	18	152	8.4
Bettiah	32	248	7.8
Bhojpur	35	269	7.7
Gopalganj	20	229	11.5
Motihari	47	305	6.5
Muzaffarpur	38	316	8.3
Nalanda	40	321	8.0
Patna	74	678	9.2
Saran	36	337	9.4
Sitamarhi	24	177	7.4
Siwan	28	240	8.6
Vaishali	27	254	9.4
Total	**419**	**3526**	**8.4**

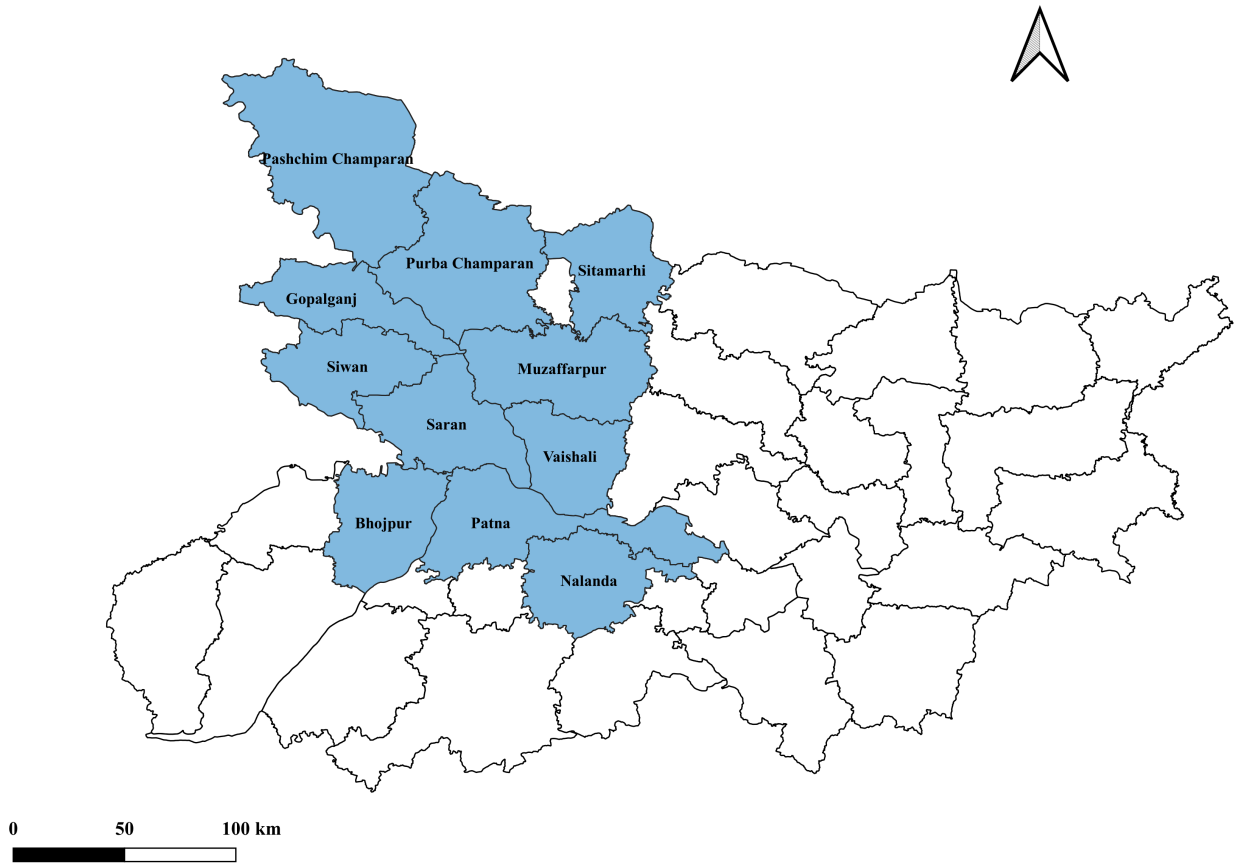


Figure 2: *Districts that were part of the study*

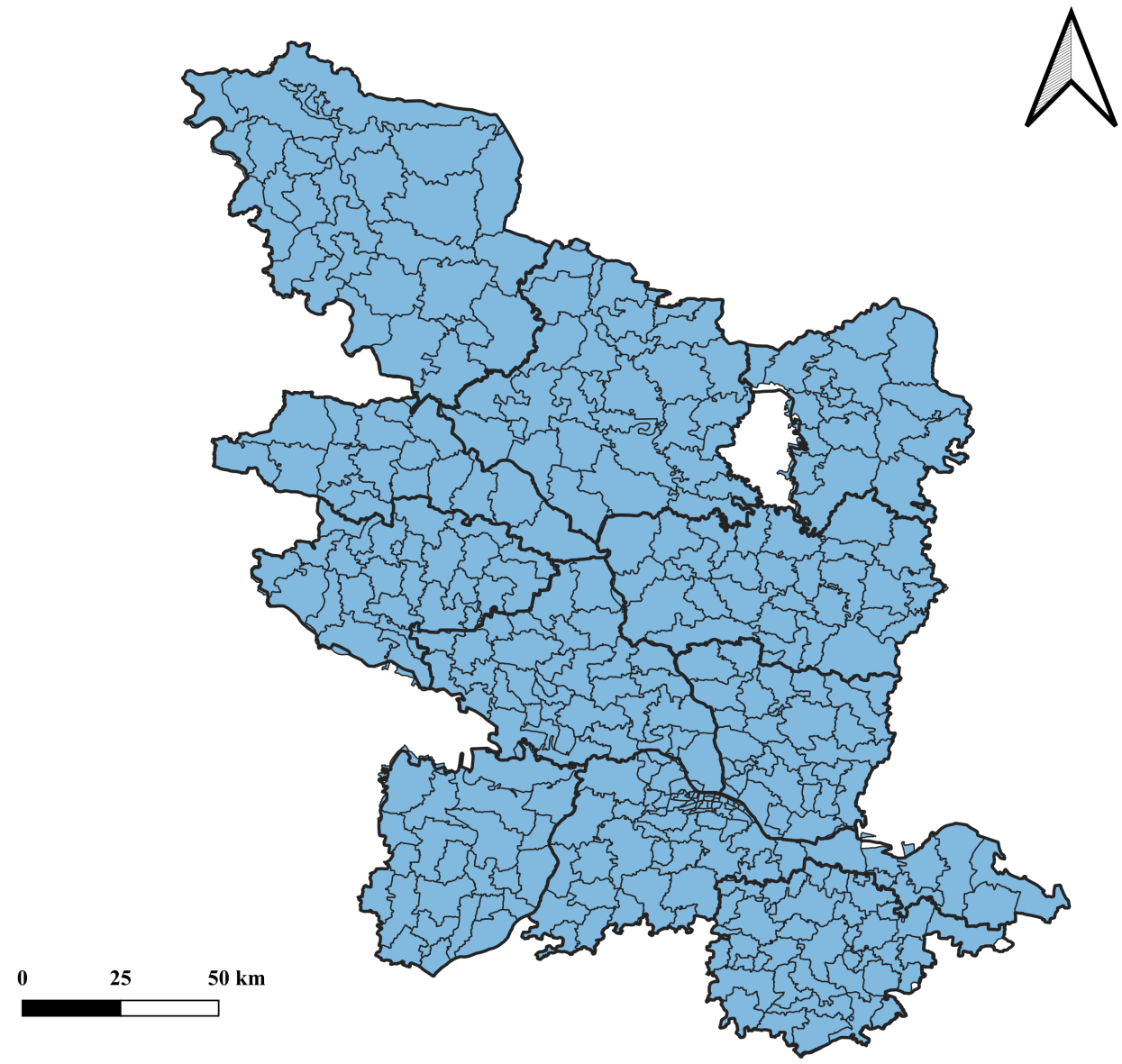


Figure 3: *Police station jurisdictions (district boundaries are marked in bold). Source: Police jurisdiction shapefile, Bihar Police*

3.2 Female Constables

For the female constables, no specific sampling method was followed as we already had the details of the police stations included in the sample. The survey team (composed exclusively of female enumerators) were tasked with going physically to the police station, asking the SHO about the female constables who were assigned duty at that specific station. Following that, the enumerators spoke directly to the female constables, obtained informed consent from them, and conducted the survey. The table below shows the district-wise distribution of female constables. On average, four female constables were assigned to each police station.

Table 2: Distribution of female constables by district

District	Police stations	Female constables	Female constables per PS
Bagaha	18	69	3.8
Bettiah	32	113	3.5
Bhojpur	35	95	2.7
Gopalganj	20	31	1.6
Motihari	47	268	5.7
Muzaffarpur	38	196	5.2
Nalanda	40	186	4.7
Patna	74	492	6.6
Saran	36	147	4.1
Sitamarhi	24	65	2.7
Siwan	28	65	2.3
Vaishali	27	83	3.1
Total	**419**	**1810**	**4.3**

3.3 Audit Exercise

For the decoy exercise, we utilized our prior interactions with the district SPs to choose the districts where we were most likely to get permission to conduct this exercise. The districts selected for this exercise were Bhojpur, Gopalganj, Muzaffarpur, Nalanda, Sitamarhi, and Siwan. Each police station in the district received three visits from the standardized victims (SVs) who were trained specifically for this activity.

Table 3: Police station visits by district (Decoy exercise)

District	Police stations	Visits per police station	Total visits
Bhojpur	35	3	105
Gopalganj	20	3	60
Muzaffarpur	38	3	114
Nalanda	39	3	117
Sitamarhi	24	3	72
Siwan	27	3	81
Total	**183**		**549**

4 The Intervention

This intervention makes several novel contributions. First, our intervention attempts to change the prevalent social norms regarding gender-based violence (GBV) to a set of norms supported by the law, explicitly connecting the gender-equal norms endorsed by the training with their identity as a 'police officer'. This is accomplished by making police officers set their aspirations and voice their grievances as a group and connect it to an ideal vision on the first day of the training. We try to subvert locally available narratives for why Violence Against Women (VAW) takes place using applied theatre, and theatre of the oppressed. For instance,

police officers we talked to during our scoping said that women who filed domestic violence complaints want to exploit GBV laws to threaten their husbands for greater control over household resources and decision-making. The intervention attempts to change mental models by re-categorizations. Our training, therefore, has sessions that deconstruct stereotypes, gender, and prevalent myths of VAW from different communities.

Finally, as our baseline data reveals, police officers have low levels of knowledge and skills relating to GBV. Even when officers know the relevant laws, the problem is compounded by the fact that women who approach the police station do not always want to formally register complaints about fear of stigma, lack of financial independence, or lack of institutions for accessing support in the aftermath. Thus, soft skills are of paramount importance in addressing women's grievances.

Our training uses embodied theatre techniques to make officers practice soft skills while interacting with them. Furthermore, we complement this with a handbook that provides a detailed protocol to follow on GBV cases.

4.1 Ideation

This section describes the process through which the curriculum was framed and an implementation plan designed. First, we consulted with former and active police personnel to identify the gaps in dealing with cases of GBV for the police which are not adequately addressed through mandatory training. Second, we then spoke with domain experts and training practitioners to identify a suitable format in which the bespoke training could be administered effectively. Finally, we conducted field visits to police stations and spoke to active officers on the ground to gauge the intensity at which the training should be administered to make a tangible impact.

Following field visits and consultations with former police personnel, including senior Indian Police Service (IPS) officials, we identified the prevalent gaps as follows.

1. bias and prejudices against women in particular
2. inability to respond with empathy
3. patriarchal notions leading to victim blaming
4. non-clarity of issues around consent
5. a general lack of agency

Informed by the on-ground teams, it was assessed that the knowledge gap was relatively less and easy to bridge, but the key challenge was in the preconceived notions of gender bias from which the attitude and behaviors of the officers were being influenced. With this knowledge in mind, the team went to brainstorm on the specific modules that the training would involve that would adequately address these gaps to empower trained officers to better handle cases of GBV. With this clear focus, key themes were identified and session objectives framed around them. The final objectives towards which the training would build were framed as:

1. Sensitizing the participants on issues of gender and gender-based violence
2. Building emphasis on the need, value, and attributes of service-centric policing
3. Improving the quality of police interaction with survivors of gender-based violence

4.2 Designing the Training Curriculum

The training curriculum applies a pedagogical technique based on Applied Theater and Theater of the Oppressed techniques, based on the pioneering works such as Boal (2014) Freire (2020) and Blatner and Blatner (1988), which utilize immersive and experiential learning techniques which have been found to be effective in different contexts in various studies. In their paper, Hoff et al. (2021) use an applied theatre method to revise norms about wife-beating and document a reduced occurrence of domestic violence in villages of West Bengal. Heller et al. (2017) use similar skill-building and immersive experience techniques as our curriculum to find a reduction in crime among economically disadvantaged youth in Chicago.

Our curriculum was developed in the year 2019 in collaboration with *'Mittika'*, a Delhi-based NGO and India's leading expert in using applied theatre and participatory approaches to drive social change on topics of gender, caste, and other forms of social inequalities. It took several months since inception to arrive at the final version of the curriculum. Several high-ranked IPS officers were also actively involved in the design of the curriculum. There was also an active involvement of the research team who provided inputs based on the field understanding of the police through qualitative scoping and police officer semi-structured interviews. Lastly, there were critical inputs from our lead trainer, Ms. Jaya Iyer,[1] who also contributed in making our training curriculum for the Police.

The discussion of the results culminated in concrete training curriculum (detailed below) as well as a handbook on dealing with GBV that was to be circulated among the officers who participated in the training. Both Hindi and English versions of this handbook were available for the officers to consult freely.

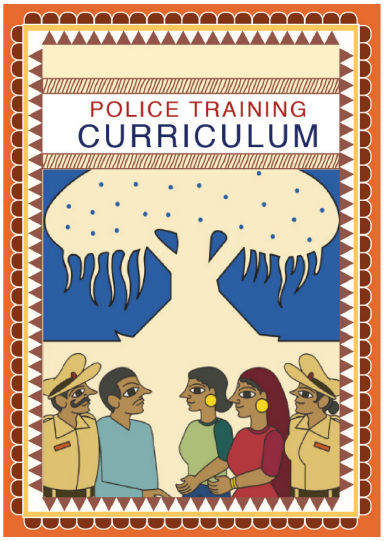


Figure 4: The GBV training handbook

4.3 Designing the Training Workshop

Since the training curriculum require active engagement, reflection, and articulation by every participant it was suggested that the overall design be experiential with enough time and personal reflection. The training design team decided that the workshop should have a maximum of **25 police officers** in a group with **three full days** (approximately 24 hours) of training. Most importantly, the workshop design had to ensure that it:

1. Creates a safe environment where dissent and disagreements are valued rather than discarded.
2. Adheres to a process that has enough time and space for peer discussions and hearing different voices, and provides ease and fun
3. Is experiential and helps to step into the shoes of others – primarily survivors of GBV

Without these aspects, the module can become top-down, and information-centric thus not serving its purpose. So, a graded step-by-step approach was built to address denial and resistance regarding GBV.

- **Day one - setting the context and preparing the ground**
  - Begins with some fun activities and games. Creating a non-threatening and open environment was important for any real work to happen

[1] An award winning theater professional and an applied theater practitioner with three decades of experience. She is also the recipient of the prestigious UNESCO Aschberg Bursary to study the Theatre of the Oppressed with Augusto Boal.

- Connecting to their idea of an ideal police station and from there looking at what was in their locus of control and influence and differentiating it from what lay beyond it was very important for enabling their sense of agency. At various junctures, the facilitator would remind them about their locus of control and influence
- The purpose of policing ensured that the focus on service-centred policing and women as half of the population was centre staged
- A fun but engaging session to accept that we all carry prejudices and it can impact our attitudes and behaviour was the last
- No session was done on GBV per se on this day but only the ground was to be prepared so that the hard-hitting content of the second day could be absorbed

Table 4: Activities on the first day of the intervention

Day - 1			
Session No.	**Session Name**	**Objective**	**Reasoning**
Session 1.1	Introduction	– Make officers feel comfortable – Set the premises for the workshop – Align expectations	Lays the groundwork for trust and open engagement
Session 1.2	Envisioning the Ideal	– Articulate the current realities of the officers as individuals and as a collective (police person and the police force) – Enlist attributes of an 'ideal' and build a case for working towards this	Helps motivate participants by imagining a better version of their role and context
Session 1.3	Purpose of Policing	– Build an understanding of the constituency – Analyze the purpose of policing	Connects policing to community service and democratic accountability
Session 1.4	Of Perceptions and Bias	– Examine concepts of bias and stereotypes – Analyze the impacts of bias and stereotypes on police duties	Introduces self-awareness and promotes impartiality in policing

- **Day two – challenging beliefs and addressing the core theme**
  - A session on inequality and discrimination helps map the reality of access in general and how some sections of society may have no access to justice
  - Understanding gender and gender-based stereotypes is a hard bite into the issue and hand and is followed by the cycle of violence
  - This flows into victim blaming – which is when for the first time we talk about a survivor of gender-based violence

– At the end of the day, some open time is suggested as the day is heavy and cognitively challenging. Singing and some light-hearted banter is needed for release and recovery

Table 5: Activities on the second day of the intervention

Day - 2			
Session No.	**Session Name**	**Objective**	**Reasoning**
Session 2.1	Recap	– Recap of the previous day	Reinforces key takeaways and prepares participants for new content
Session 2.2	Of inequality and discrimination	– Articulate linkages between power, inequality, and gender	Encourages reflection on structural factors behind gender-based discrimination
Session 2.3	Gender Stereotypes and Violence	– Analyze the impact of gender stereotypes	Helps identify how stereotypes normalize or enable violence
Session 2.4	Life Cycle of Violence	– Consolidate the different forms of GBV at different stages of women's lives	Illustrates how GBV is systemic and persistent throughout life stages
Session 2.5	Victim Blaming	– Explore who is responsible and accountable for acts of GBV – Discuss the meaning and implications of victim blaming	Challenges harmful narratives and fosters accountability in law enforcement
Session 2.6	Open Space	– Destress and get ready for the next day	Provides a space for emotional decompression and reflection

- **Day three -focus on skills and actionable takeaways**
  - The day begins with a quiz about GBV and related legislations. This is followed by a practice session on survivor-centric policing and mapping of action points when dealing with GBV cases
  - The final session is about personal action plans and feedback

Table 6: Activities on the third and final day of the intervention

Day - 3			
Session No.	**Session Name**	**Objective**	**Reasoning**
Session 3.1	Recap	– Recap of the previous day	Helps reinforce learning and maintain continuity across sessions

Day - 3 (continued)			
Session No.	**Session Name**	**Objective**	**Reasoning**
Session 3.2	Of Rules and Laws	– Examine technical knowledge of laws pertaining to GBV – Explore the manual	Equips officers with legal clarity and reference materials for GBV cases
Session 3.3	Survivor-centric Policing	– Examine police interaction with victims/survivors – Develop skills to interact with victims/survivors	Promotes empathetic, rights-based approaches in victim/survivor interactions
Session 3.4	Mapping Action Points	– Examine and improve police interaction – Articulate action points to handle GBV cases better	Encourages concrete commitments and reflection on policing practices
Session 3.5	Change Ahead	– Highlight key learnings from the workshop – Create an action plan	Solidifies key messages and supports behavior change through planning

4.4 Developing a team of facilitators

The screening and selection of the facilitators who would be administering the training looked at the following attributes:

1. Clarity and conviction regarding gender equality and justice
2. Stamina and skills for facilitating experiential processes
3. Experience in working with people in authority was desirable
4. Facilitation acumen, trouble shooting and improvisational skills
5. Resilience, teamwork and self-motivation

The Training of Trainers (ToT) was organized in September, 2022, which included a six-day intensive training and dry-runs of the Police Training Curriculum. Our lead trainers for the training were Ms. Jaya Iyer and Ms. Sutapa Sanyal (former Director General of Police, Uttar Pradesh). There were a total of 19 facilitators who were invited to be a part of the ToT, out which 8 Trainers were a native of Bihar and the rest were from other states of India. There were a total of 10 female and 9 male facilitators and they all had considerable experience of working as an applied theater practitioner. The six-day ToT included intensive reading sessions of the curriculum, pair-wise dry runs of each session included in our training curriculum followed by a detailed peer and participant feedback sessions.

At the end of each training day, there was a detailed reflection session which consisted of a detailed summary of the entire day followed by another round of peer feedback and a detailed individualized feedback shared by the lead trainers. The ToT was followed by a 3-day pilot training session organized in Arwal district from September 18-20, 2022, which was attended by 17 officers. Throughout the ToT and the pilot phase, the facilitators were consistently graded by the lead trainers on the basis of several parameters such

as understanding of GBV and related issues, content delivery, energy and conviction, teamwork and process flow. At the conclusion of the ToT and pilot, 14 out of 19 trainers, i.e. about 74% of the facilitators were retained for the main intervention.

Based on this process and the resulting feedback, it was decided that each batch would be assigned a pair of facilitators to distribute responsibilities and reduce the psychological strain of conducting the sessions.

4.5 Implementation of the intervention

Prior to the roll-out of the training, we had to follow some important steps in order to ensure the smooth implementation of the intervention. This was important because of the high case-load of police officers in the sample stations and prevailing timetables of duty, which needed to be adhered to. We had to come up with a standard operating procedure (SoP) that we could follow in all the districts in the sample to ensure maximum attendance among the officers who had been selected to participate in the training.

1. **Staggered implementation:** After conducting the pilot, the research team came to the conclusion that rolling out the intervention simultaneously would stretch the resources of the implementing partner too thin. Additionally, monitoring of the intervention was crucial and the research team lacked the personnel to do this effectively if training was conducted in multiple districts simultaneously. Therefore, the decision was made to stagger the implementation of the training with only one or two districts having the training at a single time.

2. **Detailed discussions with high level police officials:** Detailed discussions were undertaken with the high level police officials, which include senior IPS Officials at the Director General (DG) and Additional Director General (ADG) levels in the Bihar Police. The discussions were firstly to ensure that our curriculum and our intervention design was well aligned with the context of the Bihar Police. Secondly, given that our implementation was staggered, the high level officials also intimated us about the feasibility of the implementation of our data collection activities as well as training intervention at the district level - this was done to avoid any clash of our research activities with any important law and order situations wherein the police forces might be stretched.

3. **Discussions with the Superintendent of Police (SP):** The next and perhaps one of the most critical steps for ensuring full cooperation in the data collection activities as well as attendance in the training intervention was to get a full approval from the district SP, who was usually the nodal officer in chief, responsible for managing all the law and order situations for a particular district. Once the research design and the curriculum objectives were thoroughly explained to the SP, a letter of approval was issued on his behalf to all the selected police stations to ensure full cooperation. The SP also assigned a nodal officer (usually a Deputy SP) who was responsible for coordinating all our activities on their behalf. Lastly, in order to avoid any contamination of our sample, it was made sure by our nodal officer that only the selected police stations (PS) were assigned the treatment. We were also provided with the updated list of police officers at the PS level to enable us to track all the intra and inter district transfers which occurred during the course of our studies. The SP and the nodal officer also helped us to finalize the batches from the treatment police stations to make sure that the business as usual in a police station was minimally affected and we did not end up including officers in our treatment list who were not in our sample or were transferred/in long-leave/in emergency deputation.

4. **Conducting the baseline:** It was important that we conducted the baseline survey in all the police stations in a specific district before training could be commenced. This was a crucial step in order to enable a rigorous evaluation of the impact of the intervention. Additionally, this provided us with a list of officers that we could crosscheck with the list provided by the office of the SP and ensure their attendance in the training.

5. **Preparation of the field team:** A field team consisting of enumerators (for administering the reflection surveys), Project Supervisors (for ensuring all the logistics and attendance), Project Assistants (overall coordination) was deployed for each training session. The entire field staff reported to a Research Associate (RA), who was responsible for monitoring the entire intervention. Prior to the start of each individual training session, the Station House officers (SHOs), who were the nodal officers in

charge of the police stations, were contacted beforehand by our Project Supervisors to make sure that all the officers included in a particular batch were released from other duties during the training hours (unless emergency) to ensure maximum attendance. In case some officers from a particular batch were not available, we requested the SHO to allow the officers from the following batches to attend our training. In some cases, we also had to request the SP/Nodal officer to speak to the SHO on our behalf if cooperation was not ensured. Once we had approvals from all the levels, including the SHO, our field team then proceeded to call each individual officer from a particular batch to make sure that they attended the training and were aware of the timing/venue.

6. **Creation of batches:** Since we had to administer the intervention without drawing away too much manpower from the police stations, we came up with a solution to create separate batches that would be trained in each district. The idea was that not more than 3-4 officers from a single police station would be invited for training at one time. Therefore, each district had 3-4 batches of officers who were separately trained. On the other hand, we also tried to ensure that not much time lapsed between each batch that was being trained to ensure that knowledge spillovers were minimized.

4.6 Monitoring of the intervention

Following the completion of each week of training, online feedback and reflection meetings were organized by our lead trainer for 2 hours with facilitators who had facilitated the workshops in the preceding week. Our lead trainer also conducted audits of some of the sessions in order to gauge the on-field challenges and improve the facilitation process. The research team also encouraged facilitator self-reflection and a detailed peer-feedback session post the completion of each day of training and each session lasted for about 45 minutes. Lastly, the research team also maintained a facilitation monitoring document in order to make sure that the objectives of each training session were met and the facilitation flow was well aligned to our theory of change. The facilitation monitoring document can be accessed from here. The following monitoring documents were thoroughly maintained by our field staff/RAs:

1. **Detailed attendance records:** Batch-wise attendance records of each officer in the treatment were maintained thoroughly by our field staff.

2. **Field notes:** Detailed field notes (usually batch-wise) for the intervention in each district were maintained and updated regularly by the RAs. It included all the minute details about a particular batch of training, ranging from the logistics, to the performance of the facilitators, the response of the officers, the flow of our curriculum, and most importantly, qualitative records of behavioral change of the officers to GBV were maintained by our RAs.[2]

3. **Monitoring sheet:** The RAs also maintained a detailed grading sheet / monitoring document to track the progress of each training session on various parameters of a satisfactory batch of training, which included - logistical aspects, engagement of the trainers, engagement levels of the police officers, attendance, and participation. [3]

4. **Reflection sessions with facilitators:** After the completion of each day of training, it was made sure that a detailed reflection session with the facilitators and the field team and the RAs were held - this allowed for a detailed discussion around the curriculum flow, logistics, the day-to-day challenges in the implementation of the curriculum and most importantly, it could also be envisaged as a space to ensure the mental well-being of both the facilitators and the field team.

5. **Review meetings with the lead facilitator:** A monthly review meeting (usually online) was organized between our lead facilitator (Ms. Jaya Iyer) and all the facilitators to discuss the peer-feedback and overall evaluations of each of the sessions - each individual session had its own challenges, hence, such review meetings created a comfortable space for all the facilitators to share their experiences and learn from each other's shortcomings as well as improvisations depending on the situation.

[2]The field notes can be found here: https://www.dropbox.com/scl/fi/1rv6ycre3d9lmqnuu9f1m/04022023_Intervention-Notes-2022-2023_AG_SB_DB.gdoc?dl=0&rlkey=5crs632j1k2747e8cxnv8aphe

[3]The monitoring document can be found here: https://www.dropbox.com/scl/fi/5154tqeisafpi956sy2v2/02032023-GBV_Intervention-Field-Monitoring_SB_AG_DB.gsheet?dl=0&rlkey=ljfggb3svvhispe7t6rcmapxj

6. **Monitoring visits by the lead facilitator:** Our lead facilitator undertook bi-monthly field visits during the course of our intervention to make sure that all the learnings from the curriculum were implemented efficiently and the facilitators were well motivated and engaged with the training. Also, she keenly observed and noted down her perspectives on whether the officers were engaging efficiently with our training program and whether there was legitimate behavioral change which could be observed over the course of the three days. Most importantly, being an applied theater professional, she made sure that the facilitators were able to manage the facilitation process

5 The Decoy Experiment

5.1 Objective

The key objective of the Decoy Experiment was to help us evaluate whether our theater based GBV training intervention conducted in Bihar had a measurable impact on the behavior and responses of police officers when confronted with potential GBV cases. While we had already conducted a rigorous data collection exercise at the police station level across our 12 study districts, the self-reported data often has the tendency to reflect a social desirability bias. Hence, we intended to conduct a decoy experiment to supplement our data collection exercise and assess the efficacy of our training program in a business as usual setting and minimize such 'Hawthorne Effects' wherein the officers might alter their responses based on the perception of being observed. We took inspiration from Banerjee et al. (2021) for the idea of a decoy visit to observe officer behaviours in a business-as-usual context.

5.2 Sampling

The decoy experiment was conducted in six sample districts which would cover 183 Police Stations out of the total of 419 Police stations, i.e., nearly 44% of our sample. Since we had conducted the randomization district-wise and in a staggered manner, each district roughly had an equal share of treatment and control police stations with balanced observable characteristics across the two groups. The decoy visits were conducted across both treatment and control Police stations in the selected districts. There were 3 decoy visits conducted per police station in the given timeframe where the decoy visited to report one GBV and one non-GBV case respectively.

Since conducting this exercise required close cooperation and unwavering support from the Superintendent of Police (SP) of the district, we only selected those districts for the purpose of our study in which the SPs showed a keen interest. These districts were Bhojpur, Gopalganj, Muzaffarpur, Nalanda, Sitamarhi, and Siwan.

5.3 Standardized Victims (SVs)

We used Standardized Victims (SVs) to conduct the decoy visits in the selected stations following the approach of Das et al. (2016). The standardized victims were 'fake victims' who would attempt to register a case in a police station just like a regular victim.

The standardized victims were trained to register a case in a police station and answer the follow-up questions in defined patterns so that it remained consistent across the police stations. The SVs were trained to register very specific types of both GBV and non-GBV cases in a police station and were meticulously trained to answer the follow-up questions by the investigating officer in a consistent manner. The case to be presented in a particular station were randomly selected beforehand by the research team.

Immediately after the interaction (within an hour), the SVs were required to fill up a structured questionnaire that would document the quality of interaction and record the attitude of the investigating officer towards the case that was presented. The SV was also requested to answer a set of structured questions regarding the quality of the interaction and the behaviour of the officer in general, which were voice recorded.

5.4 Implementation Protocols

1. **Permissions and Approvals:** We had to seek full cooperation from the district SP before conducting the decoy visits. All the police stations in the selected districts were duly informed by a signed letter from the SP that a decoy visit by an SV was to be conducted in their police station in the coming three months. However, the timing and identity of the SV would not be revealed. In this, we followed the strategy of Banerjee et al. (2021).

2. **Timing of the Decoy:** The timing of the decoy visit in a particular police station were randomly determined by the research team. The decoy visits were conducted in different waves in order to mitigate the 'radio' effects. We expected active cross-communication across the police stations once the SV revealed their identity.

3. **Information to Police Officers:** Police officers were aware that a potential decoy experiment might take place and an official letter from the SP was addressed to every Station House Officer (SHO) in a district.

4. **Appearance of the SV:** The SV were dressed appropriately as per the local context and it was ensured that their appearance aligned very closely with that of a real victim of that particular case. We had both male and female SVs as decoys. The female SVs were accompanied by male SVs wherever appropriate.

5. **Nature of Crimes:** The crimes chosen for the decoy experiment were relatively petty and non-serious in nature. The research team trained the SVs to present one relevant GBV and one relevant non-GBV crime to the officer in charge. They were also trained extensively on how to answer the follow-up questions on that particular case so that consistency across the police stations was ensured. Assignment of the case to the police station was randomized. Most importantly, while the cases were relatively 'petty', they would still be serious enough to warrant an on-spot *'Kaccha FIR'* (or pre-FIR) or FIR registration as per the law.

6. **Recording the Interactions:** Immediately after the completion of the visit, the SVs were administered a sort of exit survey, ideally to be completed within one hour of the interaction. This survey would record the quality of their interactions with the officer in charge and the attitude of the officer towards the SV and the case that was presented. Additionally, the SVs would also record whether the police officer suspected the decoy during the interaction or not.

7. **Decoy Reveal:** In order to preclude the possibility of a fake case being registered in the official diary, the SV revealed his or her true identity in front of the officer and presented them with the official letter issued by the SP. In an alternative scenario, if the officer in charge failed to register an official report which the SV requested, a supervisor from the field team would duly inform the SHO or the officer in charge about the decoy visit.

5.5 Mitigating the Potential Risks

1. **Safety of the SVs:** To ensure the safety of the SVs, all the SHOs of the respective police stations were duly informed of a potential decoy visit within a given time frame, though the identity and timing of the SV were unknown. The research team sought full cooperation from the district SP to ensure smooth implementation. Moreover, each decoy visit was closely coordinated by a dedicated team of field supervisors.

2. **Confidentiality:** The research team ensured full confidentiality of the identity of the SV as well as the officer in charge, and all data was anonymized and stored in encrypted servers.

3. **Avoid Fake Registrations:** The SVs were duly trained to inform the police officers about the decoy visit to avoid the possibility of registering a false case. The SV would make sure that there were no official online or offline records of the false case being stored by the officer in charge.

4. **Psychological Impact on Decoy Agents:** Participating as an SV in sensitive scenarios can be emotionally challenging. The implementation partner would ensure the ready availability of psychological support for the SVs before, during, and after their interaction. Additionally, counselling services were provided.

5. **Debriefing the Police Officers:** The SVs were trained to conduct short debriefing sessions with officers after the conclusion of their interaction to address any remaining concerns. The emphasis of this debriefing session was to reiterate the study's objective as an assessment of training effectiveness rather than an evaluation of individual officers.

6. **Counselling Support:** Psychological support and counselling services were made available to SVs before, during, and after their interactions to address any emotional or mental health concerns.

References

Banerjee, A. V., Chattopadhyay, R., Duflo, E., Keniston, D., and Singh, N. (2021). Improving police performance in rajasthan, india: experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy*, 13(1):36–66.

Blatner, A. and Blatner, A. (1988). Foundations of psychodrama. *New York*.

Boal, A. (2014). Theatre of the oppressed. In *The Improvisation Studies Reader*, pages 97–104. Routledge.

Das, J., Holla, A., Mohpal, A., and Muralidharan, K. (2016). Quality and accountability in health care delivery: audit-study evidence from primary care in india. *American Economic Review*, 106(12):3765–3799.

Freire, P. (2020). Pedagogy of the oppressed. In *Toward a sociology of education*, pages 374–386. Routledge.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). Thinking, fast and slow? some field experiments to reduce crime and dropout in chicago. *The Quarterly Journal of Economics*, 132(1):1–54.

Hoff, K., Jalan, J., and Santra, S. (2021). Participatory theater empowers women.