CONFRONTING *Prejudice*: UNCOVERING STEREOTYPES AMONG POLICE OFFICERS IN INDIA *

Sofia Amaral[†] Kim Chaney[‡] Victoria Kaiser [§] Nishith Prakash[¶]

Abhilasha Sahay[‖,**]

October 21, 2024

Abstract

This paper examines the effects of a psycho-social intervention aimed at addressing plausible prejudice and bias held among police officers while handling cases of gender-based violence (GBV) in India. In collaboration with the police department at Madhya Pradesh (second largest state in India), we ran a lab-in-the-field experiment where male and female officers were randomly confronted about their mishandling of a fictitious GBV case. We find that the confrontation leads female officers to place more emphasis on the victim's account and have higher probability of pursuing a GBV complaint than officers in the control group. Male officers, on the other hand, have the counter-reaction: they lay less emphasis on the victim's account and there is no significant effect on registering a GBV complaint. A potential explanation for our findings is that - at baseline - female officers show a milder bias against GBV victims, while more than half of the male officers are strongly biased. Considering that policing is a male-dominated profession, female police officers may perceive a bias in their work environment that is stronger than their own and thereby, de-bias their case handling on being confronted. Male officers showcase a behavior consistent with backlash, likely driven by strongly biased men. This study sheds light on the role of behavioral nudges in addressing possible biases among public sector workers and demonstrates how response to such nudges could be shaped by the extent of underlying bias.

Keywords: Prejudice Confrontations, Gender-Based Violence, Police, India
JEL Codes: J12, J16, K42, C93, C91

*We are grateful to Madhya Pradesh Police and Central Academy for Police Training for the collaboration on this intervention. The views described in this paper do not necessarily reflect those of the Madhya Pradesh Police. We appreciate the valuable feedback from Prashant Bharadwaj, Michael Kaiser, Paul Niehaus and Helmut Rainer. This study received bioethics approval from the University of Connecticut under the protocol X21-0091 approved May 21, 2021. We thank Vishakha Wadhwani and Asmi Khushi for their fieldwork and their superb work as research assistants, and Anup Roy, Sitaram Mukherjee, and all the field staff at DAI Research and Advisory Services for their invaluable contributions to data collection. The study was pre-registered with the AEA Trial Registry (RCT ID: AEARCTR-0008611). All errors and omissions are our own. We appreciate funding from University of Connecticut, Joachim Herz Foundation, and Leibniz Association.

[†]World Bank. Email: s.amaral@worldbank.org

[‡]University of Connecticut. Email: kim.chaney@uconn.edu

[§]Née Endl-Geyer. ifo Institute. Email: kaiser@ifo.de.

[¶]Northeastern University, CESifo, IZA, GLO, HiCN, and CReAM. Email: nprakash@northeastern.edu

[‖]World Bank. Email: asahay@worldbank.org. (Corresponding Author)

[**]Author order determined alphabetically

1 Introduction

Gender-based violence (GBV) remains widespread in most parts of the world with low-to-middle income countries being the most affected (Sardinha et al., 2022).[1] Police response and formal reporting of GBV is critical to induce deterrence against such crimes. Yet, past evidence suggests that complaints of GBV may often be discredited or dismissed by police officers due to plausible biases, doubts around the veracity of such complaints and victim-blaming; thereby, diluting the impact of police responsiveness and more fundamentally, undermining victims' trust in institutions and willingness to report crimes against women (García-Moreno et al., 2015; Belknap, 2010). Addressing such biases, is therefore, a necessary (albeit not sufficient) condition to ensure equal access to justice and mitigate incidence of GBV.

Confronting people with their prejudices is found to be an effective tool for reducing bias and changing behavior among perpetrators, especially in contexts of racism and sexism (Alesina et al., 2018; Mallett and Wagner, 2011).[2] After a prejudice confrontation, people strive to be more egalitarian and self-regulate their bias (Chaney and Sanchez, 2018). However, this research has overwhelmingly focused on prejudice in high-income countries where cultural norms signal egalitarianism (Jayachandran, 2015). When people receive feedback that threatens their self-image, they may also engage in defensive strategies such as denying, downplaying, or dismissing the feedback (e.g. Howell et al., 2017). This defensiveness could increase peoples' bias.

Against this background, this study examines a novel intervention of prejudice confrontation in the context of GBV and Indian police officers. We investigate prejudice confrontation and its effects on individuals' bias in a setting where norms of egalitarianism are weak or nonexistent. To the best of our knowledge, this research is the first that examines the impact of prejudice confrontations at the workplace, in the context of state actors' biases around GBV. We collaborated with Madhya Pradesh Police (second larget state in India) to design and implement a randomized controlled trial (RCT) in the form of a lab-in-the-field experiment. Our sample in-

[1] Gender-based violence can include sexual, physical, mental, and economic harm, threats of violence, coercion, and manipulation. While GBV can affect people of all genders, rates of GBV against women and girls are particularly high (https://ncadv.org/STATISTICS). GBV can take place in private or in public and takes many forms that can include intimate partner violence, sexual violence, child marriage, female genital mutilation, and so-called "honor crimes" (Degener and Koster-Dreese, 1995). Reliable cross-country estimates are not available for most forms of GBV, yet, (Sardinha et al., 2022) estimate that worldwide 27 percent of women are victims of intimate-partner violence.

[2] Prejudice confrontation are defined as verbal challenges directed at those who commit a blatant, subtle, or unspoken act of discrimination (Chaney et al., 2015; Czopp et al., 2006).

clude both male and female police officers. Participating officers reviewed two cases, one GBV case and one non-GBV case, and answered a computer-based survey on how they would handle the cases. Officers were randomly assigned to a treatment condition where a high-ranking police officer confronted them about their bias in dealing with the GBV case, after the officers completed the survey. The confrontation consisted of a private in-person conversation with the senior officer. Officers in the control condition received neutral feedback on case handling. One week later, all officers reviewed a GBV case again and solved a computerized stereotyping reaction time task in which pictures of potential victims were shown and they were asked to rapidly categorize descriptions that might apply to the victim or not. The study design allows us to understand future responses to GBV crimes by police officers in two dimensions. First, whether and how confronted officers change their behavior in handling GBV cases? We hypothesize that the police officers in the treatment group will showcase less-prejudiced behavior in handling cases of GBV than officers in the control group. Second, does the change in behavior vary by gender of the police officer. Notably, we contend that female police officers are more likely than male officers to make egalitarian changes in their behavior after being confronted with their own stereotypes concerning GBV cases because of their shared gender identity with the predominantly female victims of GBV. Yet, as women make up only a small percentage of police officers in India (13%), they may feel pressured to conform to prevailing workplace norms, which could ex-ante lead to a stereotypical response[3].

We assess change in prejudiced behavior one week later in two ways. First, we examined if police officers who were confronted for bias in handling a GBV case would demonstrate less bias in handling a future GBV case (one week later) compared to police officers who were not previously confronted for their bias in handling a GBV case. Second, we sought to assess automatic stereotypes of police officers one week after they had (or had not been) confronted. Past research on prejudice confrontations overwhelmingly focuses on changes in stereotype usage (Chaney Sanchez, 2018; Czopp et al., 2006), rather than behavior. Further, past research has found that people who were confronted for using negative Black stereotypes one week prior demonstrated significantly less stereotype activation when assessing novel Black targets compared to people who were not previously confronted (Chaney Sanchez, 2018). Notably, demonstration of lower stereotype activation among confronted police officers would suggest a deeper reflection and regulation of GBV stereotypes than mere assessment of handling of

[3]National Crime Records Bureau (Ministry of Home Affairs). 2018. Crimes in India 2018 Statistic. Volume I. Government of India.

GBV cases.

Our main finding is that - while there is no statistically significant aggregate effect of the confrontation treatment - the confrontation treatment does indeed lead female and male officers to react differently: female officers respond less stereotypically in handling a GBV case and place greater emphasis on the victim's account. Female officers in the treated group prioritize the victim's statement by 23 percentage points more than the control group, which roughly corresponds to a 27% change. Confronted females also have an 8 percentage points higher probability of registering and pursuing a GBV complaint than the control group. Male officers respond in the opposite way. While there is no statistically significant effect on pursuing or dropping a complaint, male officers put less priority on the victim's statement and more on the offender's statement post-confrontation. These findings are robust to a battery of robustness checks, including multiple hypotheses testing and alternative specifications. A potential explanation for these heterogeneous findings is that majority of the female officers (72%) show mild bias in handling a GBV case before treatment. Yet, more than half of the male officers (51%) exhibit a strong bias.[4] Considering that policing is largely male-dominated, the average female officer may perceive a bias in her work environment that is more severe than her own. The confrontation treatment leads female officers to de-bias their case handling and adjust their beliefs in a professional context. However, male officers show behavior consistent with backlash that is driven by strongly biased officers. When we exclude strongly biased male officers from the analysis, we find that for our main outcome (prioritization of the victim's statement), the negative effect for male officers effectively. Taken together, the intervention has had the effect of encouraging women to de-bias their actions, which means less stereotyping when working on GBV cases. Men, on the other hand, are less receptive to the feedback. Policymakers may have to consider that a prejudice confrontation may affect female and male officers differently, based on the underlying extent of bias.

We additionally conduct a computerized stereotyping task, in which we assess officers' stereotyping when assessing a woman reporting GBV. A similar pattern of results emerges. For male officers, we find a significantly greater use of GBV stereotypes after a confrontation. That is, male officers assigned significantly more negative stereotypes to women reporting GBV when they were confronted one week prior compared to male officers who were not previ-

[4] An example of a strong bias would be the filing of a report against the GBV victim instead of the accused. We consider it a mild bias if, for example, the officer places an investigation against the victim among the top three (of 5) important investigative parts.

ously confronted. We do not find such an effect for female officers. Notably, male and female officers' assignment of negative stereotypes to men and women reporting non-GBV cases did not significantly differ by condition, demonstrating that this effect among male police officers is unique to GBV cases. This finding is consistent with our results on GBV case handling. For men, prejudice confrontation seems to trigger a backlash. Women do not change their inner beliefs, and on an automatic task, did not rely on significantly fewer stereotypes for women reporting GBV cases after being encouraged to do so by the confrontation feedback.

This study contributes to two streams of literature. The first body of work relates to the impact of confrontation-based interventions in addressing biases. Past scholarship demonstrates that being confronted on one's prejudice can reduce bias and change behavior. For example, White Americans confronted for holding negative Black stereotypes are less likely to reflect biases 1-week later (Chaney and Sanchez, 2018). Similarly, men confronted for relying on gender role stereotypes engaged in compensatory behavior (Mallett and Wagner, 2011) and reported more favorable attitudes towards women compared to men who were not confronted. Outside of lab experiments, e.g. Alesina et al. (2018) found that confronting teachers' biases against immigrants led to an increase in grades teachers assigned to immigrant students. However, this research has predominantly focused on high-income countries and on non-crime themes. With this paper, we add to the literature by investigating the effect of prejudice confrontations in the context of gender-based violence - one of the most socially sensitive topics - in a developing country.

The paper also relates to empirical research that investigates the drivers of GBV. Scholars have linked GBV to social norms (Green et al., 2020; Bandiera et al., 2020), cultural factors (Guarnieri and Rainer, 2021; Tur-Prats, 2019), labor market conditions (Aizer, 2010; Anderberg et al., 2016), liquidity constraints (Hidrobo et al., 2016), divorce legislation (Stevenson and Wolfers, 2006), and emotional cues (Card and Dahl, 2011). However, research regarding the prevention of GBV is limited. In fact, there is little evidence focusing on the role of state actors. The police and courts are arguably a necessary institution to address GBV, as their actions determine women's willingness to turn to formal support services (Palermo et al., 2014; Prakash et al., 2024), sanctioning, and affect deterrence (Amaral et al., 2023). Our work contributes to this strand of the literature by examining how addressing police bias can change police officers' discrimination and behavior toward violence against women and possibly, improve redressal of such crimes.

The remainder of the paper is structured as follows. The next section provides information about the police force we collaborated with, the participants of the experiment, randomization, and the experimental protocol. In Section 3 we discuss our empirical approach and balance tests. Section 4 reports our main results and robustness checks. Section 5 concludes.

2 Experimental Design

In this section, we describe the context, the participants of the experiment, and randomization. We then outline the procedure of the two stage lab-in-the-field experiment.

2.1 Context

The experiment was conducted in the state of Madhya Pradesh in India. This state ranks second (out of 29) in sexual assault homicides and third in dowry deaths.[5] The police play a critical role in addressing and combating GBV, particularly when considering the limited availability of specialized services to tackle it. According to the 'IPF Citizen Satisfaction Survey on SMART Policing' in India, the police in Madhya Pradesh have received notably low ratings. In the specific category 'Fair, unbiased and lawful policing' the state has been positioned at the 24th rank (out of 29), while in terms of 'Integrity and corruption free service' it has reached the 27th rank.[6] The civil police in Madhya Pradesh consists of roughly 9,000 police officers in the ranks of Sub-inspector (SI) and Assistant Sub-inspector (ASI). These officers typically handle crime reports at the early stages of investigations. About 13% of them are female.[7] The police is a very hierarchical organization and senior officers are held in high esteem by their junior colleagues. Thus, a confrontation with one's own bias by a senior in a police environment can be particularly effective.

2.2 Participants

A total of 323 officers were recruited to complete a two-part experiment.[8] Officers were invited from the districts of Bhopal, Harda, Hoshangabad, Raisen, Sehore, and Vidisha in Madhya Pradesh, India. Officers were recruited via a two-stage process. First, each station head received a communication informing them about the study and were requested to provide a list of officers eligible to participate. Second, eligible officers were invited directly via a phone call. Only officers who were available to attend sessions were ultimately included in the study sam-

[5] National Crime Records Bureau (Ministry of Home Affairs). 2018. Crimes in India 2018 Statistic. Volume I. Government of India.

[6] IPF smart policing survey 2021.

[7] Bureau of Police Research and Development (Ministry of Home Affairs). Data on Police Organization 2021.

[8] The recruited officers were either of the rank Sub-inspector (SI) or Assistant Sub-inspector (ASI). These types of officers were selected since they would typically handle crimes - including GBV reports- at the early stages of investigations.

ple. Our approach involved inviting on average two officers per station in order to minimize spillover concerns between treated and control officers.[9] The participants ultimately consisted of 58 (20%) female and 239 (80%) male officers in an age range of 24 to 61 years (mean age: 42). Participants were provided lunch on site and officers' time spent on the research was considered on-duty time.

2.3 Randomization

The experiment took place in November and December 2021. The selection of police officers to participate in the experimental days was contingent upon their scheduling availability. Two parallel sessions were held per day. The invited police officers were randomly divided into two groups on site (this division was achieved by instructing them to form a line and alternate between saying the numbers 1 and 2). The formation of these two groups was for purely organizational reasons since the computer rooms (see next) were not large enough for one large group. Subsequently, the lab-in-the-field experiment took place with these two groups, each occupying a separate computer room. The process of individually assigning participants to either the treatment or control group was integrated into the survey software and executed automatically on the computers where the officers sat (see Figure 1).

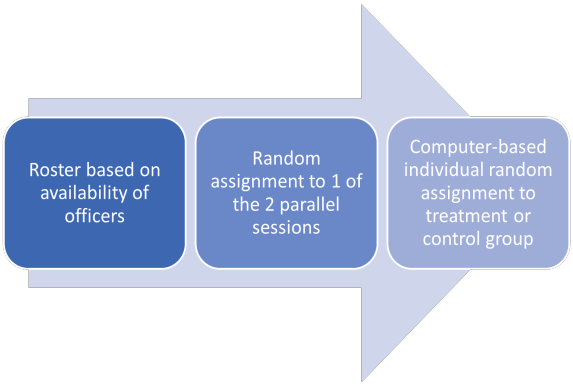


Figure 1: Randomization of participants

[9]We approached 360 officers directly via phone, thus, we achieved a response rate of 90 percent. Officers' unavailability was mostly due to being on duty, or unexpected local events that were of higher priority. For this reason, we expect that there is minimal difference between officers taking part in the experiment and those who did not.

2.4 Procedure

Baseline

The experiment relies on a between subjects design. First, after the participants had given their consent, they reviewed one neutral, non-GBV, case followed by one GBV case in a computer-based survey format. Participants completed several questions involving the cases. Starting with whether they think a crime has been committed, and then how they would proceed in the investigation step by step. The instrument included a mix of questions; there were yes and no questions, and questions where the officers were asked to rank different investigative actions in order of priority and questions that required a single or multiple choice answers. Based on these case-related questions, we constructed our key outcome variables: whether the crime is detected, whether the officers decide to officially register the complaint, whether they follow up or drop the complaint, and whether and how they rank different parts of the investigation that result from prioritizing statements from the victim or the accused. These questions were asked for two types of cases, i.e., gender-neutral, property crimes and GBV cases. We included two GBV cases, which were randomly assigned to subjects, to ensure the generalizability of findings following a confrontation. Thus, while one GBV case involved domestic violence, the second included a case involving dowry. The random assignment of one of the two GBV cases was built into the survey software and automatically executed on the computer systems. All cases were developed by the research team and were closely aligned with real cases reported to the police in this region. All cases included multiple people and claims in order to mirror the complicated, multi-party nature of real cases. They were extensively pilot-tested in separate samples of officers and iterated based on feedback from the pilot before the roll-out of the experiment.

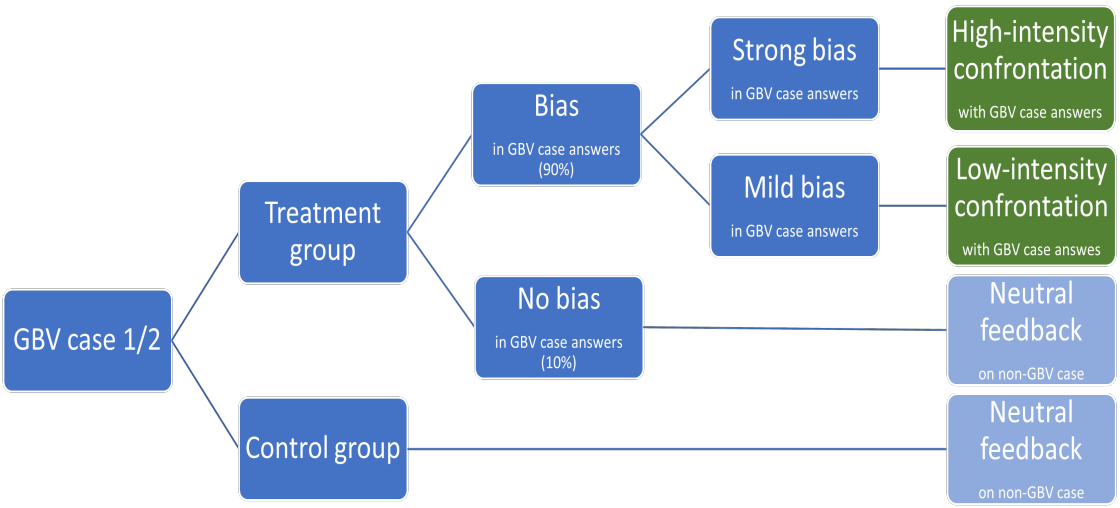


Figure 2: Procedure at Baseline

Second, after the officers completed the two cases (one property crime and one GBV case), they notified the field staff that they had completed the initial computer-based survey and waited to enter a 1-1 session with a senior officer in a separate room.[10] When called into the room with the senior officer, participants received feedback based on their randomly assigned condition. Officers in the control condition and those in the treatment condition who demonstrated no bias (10%) in their handling of the GBV case received neutral feedback unrelated to GBV (see Figure 2).[11] In contrast, participants who were randomly assigned to the treatment condition and demonstrated bias in their handling of the GBV case received confrontation feedback. This confrontation was adapted slightly depending on which of the two GBV cases they completed and also depended on the extent of the bias shown in their answers. Officers demonstrating a strong bias received a high-intensity confrontation and officers showcasing a mild bias

[10]We recruited four male senior officers to deliver the treatment. This choice was motivated by two factors. First, the experiment was intended to mimic officer's interactions while on duty. Second, in our context, due to historical factors, the police is a reserved institution. As a result, having external parties or researchers conduct the confrontation would be very challenging and likely ineffective. Since the police is a very hierarchical organization, senior officers are highly respected by their junior peers. Senior officers were recruited in the same fashion as the SI and ASI officers (via formal letter and phone calls), and also by proximity to the location of the experiment. Senior officers were not the direct managers and supervisors of the SI and ASI officers who participated in the experiment. This allows us to mitigate any potential backlash since officers operate in different jurisdictions.

[11]In a neutral feedback session, the senior officer informed the participants that in general their way of handling the cases was fine but that they 'did not pay proper attention and time to the case while answering' and '[...]when you solve a case on the computer, it is important to be careful while responding'.

were given a low-intensity confrontation. [12] Senior officers were instructed to deliver all feedback and confrontations in a neutral and non-aggressive fashion, focusing on an educational approach. Participants were given a chance to respond to this feedback, though the interaction time was cut off by field staff after two minutes. The feedback and response were audio recorded, and a member of the research team was present in the room during this interaction.

Third, after the feedback session, participants were directed back to a computer to complete a brief survey which included measures of negative feelings toward oneself (i.e., guilt/shame or feeling offended) or the senior officer (i.e., anger) ('negative self- and other-directed affect'; Chaney and Sanchez 2018; Czopp et al. 2006), as well as measures of their respect for the senior officer, how they believe they performed, and demographic information. Participants were then dismissed from the session and informed they would return in 1 week.

[12] An officer is considered to have a strong bias is he/she demonstrates a leaning towards filing a report against the victim (instead of the accused). An officer is considered to have a mild bias if, for example, the officer places an investigation against the victim among the top three (of 5) important investigative parts. If officers received the high-intensity confrontation, the senior officer stated '[...] it looks like you sided with the defendant's mother even though somebody else is the victim who came up with the complaint of domestic abuse. It seems like your actions were influenced by biased beliefs against women in society. Police officers cannot do their job based on their assumptions [...]'. Participants who demonstrated more mild bias during their response to the GBV case received the low-intensity confrontation, during which the senior officer stated: '[...] during the investigation you prioritized the statements of the defendant and his friends [...]. It seems that in this case your action was influenced by the prevailing beliefs against women [...]. Police officers cannot do their job based on their perceptions.'

Endline

One week later, participants returned to complete a computer-based survey and a measure of automatic stereotyping. The attrition rate was 8%, and did not result in a selected sample.[13] The survey included a novel GBV case (again, developed by the research team) with questions regarding how participants would handle this case. After, participants completed a measure of reflection about their handling of GBV cases, perceived truth of GBV complaints, interest in future training, and several secondary outcomes (perceived norms about GBV case handling, attitude towards women, and empathy).

After completing the survey section, participants completed a computerized stereotyping reaction time task. In this task, participants were presented with various pictures of male and female faces in the center of the computer screen, below which was a sentence representing a report to the police, such as 'This woman reported molestation by a group of boys at the bus stop near her college'.[14] For each case, after the head shot of the victim and complaint, which were shown for 8 seconds, nine words appeared on the screen above the image, and participants had the task of categorizing these words as 'Applies to this person' or 'Does not apply to this person' by pressing the corresponding key. These categories were presented at the top of the screen as a reminder for participants. Each word appeared on the screen for 2.5 seconds or until a response was recorded. The nine words were randomly pulled from a list of six victim-blaming stereotype words (e.g., liar, at fault, manipulative) and six innocent words (e.g., innocent, good, moral). Note, participants first completed two practice trials and then completed 40 test trials. Of the 40 trials, 8 included male targets reporting non-GBV complaints, 16 women targets reporting non-GBV complaints, and 16 women targets reporting GBV complaints. The two outcomes constructed based on this stereotyping task are counts of clicks that demonstrate a denial of stereotypes in the GBV and non-GBV context, respectively (i.e. how often innocent words were selected for victim images).

[13] The reason for absence was work-related unavailability and the characteristics of officers that dropped out do not differ from those of the entire sample.

[14] We modified a stereotype inhibition task (adapted from Chaney and Sanchez 2018). The task was administered on Inquisit software by Millisecond. Images were taken from the Chicago Face Database-India.

3 Empirical approach

The effect of interest is the causal impact of prejudice confrontation on police bias. We estimate the following linear model:

$$Y_{i,s,o} = \beta_1 + \beta_2 Treatment_i + \mathbf{X}_i + \gamma_s + \alpha_o + \epsilon_{i,s,o} (1)$$

where Y is an outcome of interest (e.g., priority given to victim statement, registration of complaint) for officer *i*. $\mathbf{X}_i$ is a vector of individual level controls including gender, age and post of the officer. We also include baseline session (γ_s) and senior officer (α_o) fixed effects to account for any observed or unobserved differences, based on how sessions and confrontations were conducted. The main explanatory variable *Treatment* identifies the experimental condition for officers randomly assigned to the confrontation group (i.e. confrontation with GBV case handling) in comparison to those that did not. We cluster standard errors at the officer group (session) level.

We check for balance on an array of baseline characteristics. Among the 11 indicators evaluated, there are no systematic differences between the treatment and control group officers for the majority of indicators - both within and across gender groups (Table A1). On examining baseline characteristics among male officers in the treatment and control group, we find an imbalance in 2 of the 11 variables: Officers in the treatment group prioritize the victim's statement less often than the control group and take longer to respond in the non-GBV case (Column 5). The baseline characteristics of female officers in the treatment and control group are balanced, except one (Column 6). Treated females prioritize the victim's statement in one of the GBV cases slightly less than females in the control group. In Column (7), we look at the difference between male and female officers in the treatment group and find a significant difference in officers' age and response time in the non-GBV case. Female officers in the treated group were significantly younger and males had a longer response time in the non-GBV case.

Our results are robust to the inclusion of the above baseline controls that show statistically significant differences between treatment and control groups (Table A6). Our preferred specification controls for officer characteristics, i.e., officer age and whether the officer was posted in the capital city of Bhopal. Further, we conduct a randomization inference test, closely following (Heß, 2017). If the confrontation had not affected any officer's case handling at all, then the

observed outcome for our main variable (i.e. prioritizing the victim's statement) would have occurred with a probability of below 0.2%, as measured by the p-value implied by Figure A1. We interpret this as corroborating evidence that the confrontation did in fact influence officers' case handling.

4 Results

This section first focuses on our main results. In the next step, we show robustness checks.

4.1 Primary Effects

First, we estimate the impact of the intervention on six key outcomes that capture the handling of a GBV case one week after the confrontation. The first four outcomes are variables that are constructed from survey questions on how the officers would handle the presented GBV case. They include binary variables on whether the officer recognizes that a crime has been committed, decides to drop the complaint and not to pursue the complaint (after having it registered), prioritizes the victim's statement over other aspects of the case, or prioritizes the offender's statement in the case processing. In the other two outcomes, we look at officers' general disposition toward GBV. From the questions on how they would handle the presented cases, we derive the dummy variable whether they are likely to file a complaint against the accused. From a case-independent survey that the officers completed at the end of the experiment, the outcome of how many out of 10 rape reports officers believe to be true is constructed.Second, in the computerized stereotyping reaction time task, we analyze to what extent participants are biased based on the appearance of victims. The two sub-outcomes derived from this task are counts of keyboard clicks that represent a denial of victim stereotypes (i.e. select innocent words for the victim image) in the context of GBV and non-GBV cases, respectively.

Over 95% of the officers recognize that a crime has occurred after reviewing the GBV case and there is no impact of the confrontation intervention on this outcome (see Table A2). For outcomes concerning further investigation of the GBV case, we find no impact of the confrontation - on likelihood of dropping a complaint, prioritizing victim's or offender's statement - for the full sample (Panel A, Table 1). However, this average effect masks heterogeneous effects by gender (Panel B, Table 1). Treated female officers have an 8 percentage points higher probabil-

ity of pursuing a GBV complaint than the control group (Columns 1 and 2). Further, treated female officers prioritize the victim's statement by 23 percentage points more than the control group, which roughly corresponds to a 27% change. Consistently, we observe a negative effect on the prioritization of the defendant's statement for female officers (Columns 5 and 6). For male officers, there is no statistically significant effect for dropping a complaint. However, they seem to push back after a confrontation, putting less priority on the victim's statement and more on the offender's statement. In Columns (1), (3), and (5) of Table 1 we implement fixed effects for the senior officer that confronted the officer and the experimental session. Controls for officer characteristics, i.e. the officer's age and posting, are then added in Columns (2), (4), and (6). The results for the two different specifications are almost identical in terms of effect sizes and precision.

Table 1: The effects of confrontation on the handling of GBV cases

	Drop complaint		Prio on victim's statement		Prio on offender's statement	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Without gender heterogeneity						
GBV treatment	0.011	0.013	-0.058	-0.055	0.047	0.043
	(0.030)	(0.030)	(0.050)	(0.049)	(0.041)	(0.041)
	[0.377]	[0.652]	[0.164]	[0.260]	[0.304]	[0.313]
Controls	no	yes	no	yes	no	yes
Control Group Mean	0.081	0.081	0.844	0.844	0.075	0.075
Rsq	0.052	0.069	0.038	0.053	0.064	0.072
N	297	297	297	297	297	297
Panel B: With gender heterogeneity						
GBV treatment	0.034	0.035	-0.129**	-0.127**	0.096*	0.092*
	(0.034)	(0.034)	(0.055)	(0.055)	(0.045)	(0.046)
	[0.197]	[0.197]	[0.022]	[0.022]	[0.048]	[0.048]
treatment x female	-0.115*	-0.113*	0.361***	0.360***	-0.246***	-0.246***
	(0.057)	(0.055)	(0.072)	(0.075)	(0.050)	(0.052)
	[0.094]	[0.094]	[0.001]	[0.001]	[0.001]	[0.001]
female officer	-0.022	0.015	-0.092	-0.141**	0.113**	0.126***
	(0.048)	(0.050)	(0.066)	(0.063)	(0.041)	(0.040)
Controls	no	yes	no	yes	no	yes
Control Group Mean	0.081	0.081	0.844	0.844	0.075	0.075
Rsq	0.065	0.076	0.073	0.084	0.091	0.098
N	297	297	297	297	297	297

Notes: This Table reports estimates of the effect of the intervention on the likelihood of dropping a GBV complaint (Columns 1 and 2), the likelihood of prioritizing the victim's statement (Columns 3 and 4), and prioritizing the offender's statement (Columns 5 and 6). All three outcomes are dummies and take the value 1 if the complaint is not pursued, the priority is given to the victim's statement, and the priority is given to the offender's statement, respectively. 'GBV treatment' is an indicator for being confronted with one's own bias in the GBV case handling. 'treatment x female' is an interaction term of this indicator and the gender dummy (=1 if female). The specification in Columns (1), (3), and (5) includes senior officer and session fixed effects. In Columns (2), (4), and (6) we add officer characteristics as controls (officer's age and a dummy for posting is in the capital Bhopal). Standard errors are clustered at the session level and are reported in parentheses. Bootstrapped p values of multiple hypotheses tests are shown in square brackets. We use the Stata command mhtreg for multiple hypotheses testing that provides a procedure that asymptotically controls family-wise error rate and is asymptotically balanced (Steinmayer, 2020). It is based on List et al. (2019) but modified to be used in a multivariate regression setting.

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

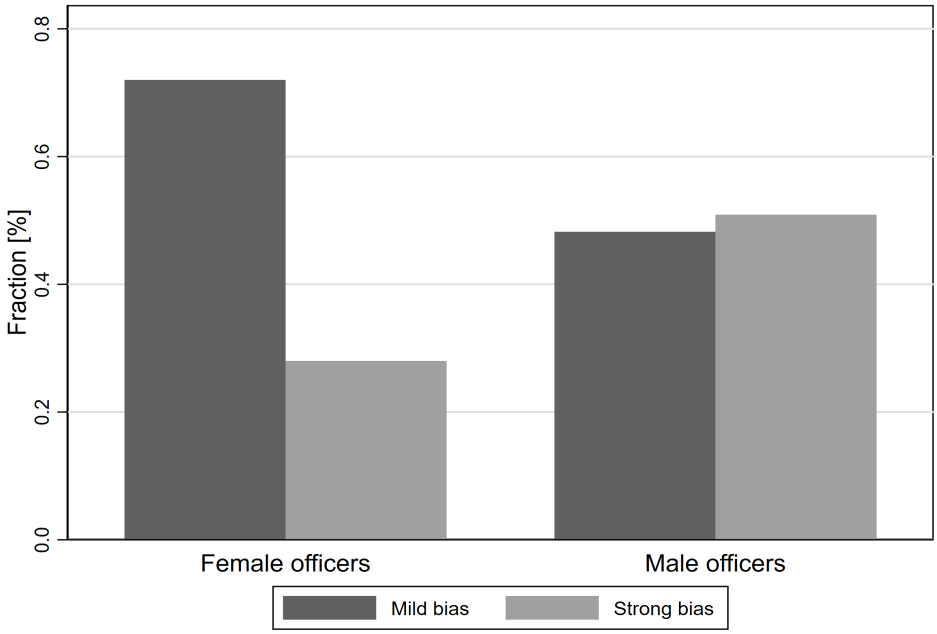
For the second set of results, which concerns the officers' general disposition toward GBV, we do not find any statistically significant impact of the confrontation. There is no differ-

ence between the treatment and control group regarding the belief in the truth of rape reports (Columns 1 and 2 of Table A3) and the likelihood of registering a GBV complaint (Columns 3 and 4). Notably, in most instances, GBV cases are registered by officers, i.e. 93 percent of cases in the control group were registered. This is likely due to the fear of administrative sanctions or penalties that officials may face if they do not comply with the requirement to register any complaint. In the same fashion as in the previous outcome table, we show a specification with senior officer and session fixed effects first (Columns 1 and 3) and add officer characteristics as controls in Columns (3) and (4). Our preferred specification includes officer characteristics controls and is used in the further course of the paper.

Next, we report results for pre-specified outcomes for which standard errors are too large to be able to draw any conclusions. The participants completed a survey on how they felt after the feedback session with the senior officer at baseline. We do not observe any statistically significant effect of the confrontation on feeling guilty or other self-reflection variables, except for 'regret' where the confrontation leads to a decrease for males (see Table A7). At the end of the experiment (endline), the officers completed another short survey on victim-blaming (in how many out 10 GBV cases they think it is the woman's fault) and their perceived norms (to what extent they think their friends, family and partner think GBV is the woman's fault). There is also no statistically significant treatment effect for those outcomes (see Table A8).

We now turn to discussing possible mechanisms for the observed result patterns. While the majority of female officers (72%) show a mild bias in handling a GBV case before treatment, for more than half of the male officers (51%) a strong bias comes to light (Figure 3). Given that policing is highly male-dominated, the average female officer perceives a bias in her work environment that is more severe than her own.[15] The confrontation treatment thus de-biases female officers and makes them adjust their professional beliefs. Male officers seemingly behave in a way consistent with a backlash, which is driven by strongly biased men. In Table A4, we exclude male officers that showed a strong bias at baseline and find that for our main outcome (prioritization of victim's statement), the negative effect for males becomes substantially smaller and statistically insignificant.

[15] Among ASI and SI officers in Madhya Pradesh, 13% are female. Bureau of Police Research and Development (Ministry of Home Affairs). Data on Police Organization 2021.



Notes: This Figure shows the fraction of female and male officers that show a mild or a strong bias in the handling of the GBV case at baseline, respectively. An example of a strong bias would be the filing of a report against the victim. We consider it a mild bias if, for example, the officer places an investigation against the victim among the top three (of 5) important investigative parts.

Figure 3: The extent of bias by gender of officer

A similar result pattern emerges in the negative stimuli reaction task where we evaluate to what extent participants are biased based on the appearance of victims. Participants solve a computerized stereotyping reaction time task in which pictures of potential victims are shown and they are asked to rapidly categorize descriptions (e.g. 'innocent', 'at fault') that might apply to the victim or not via key presses. We find that treated male participants show more agreement with GBV stereotypes in GBV trials. This indicates significantly greater use of GBV stereotypes after a confrontation (Column 1 of Table 2). On the other hand, female officers demonstrate lesser acceptance of GBV stereotypes; however, this effect is not statistically different from zero. As expected, there is no impact of the confrontation on bias in gender-neutral cases (Column 2 of Table 2). A potential explanation for the different responses by gender is consistent with the one outlined above for the case handling in Table 1. Men's counter-reaction and retaliation after a confrontation can explain this patter and women do not change their inner beliefs, but de-bias their case handling after being encouraged to do so by the confrontation feedback.

Table 2: The effects of confrontation on stereotypes based on victim's appearance

	Neg stimuli-GBV cases	Neg stimuli-other cases
	(1)	(2)
GBV treatment	-6.914**	0.676
	(2.512)	(1.242)
treatment x female	3.157	-5.373
	(4.051)	(3.180)
female officer	1.725	1.987
	(3.303)	(2.300)
Controls	yes	yes
Control Group Mean	44.828	45.867
Rsq	0.235	0.253
N	233	233

Notes: This Table reports estimates of a computerized negative stimuli reaction task. The outcome is the number of key presses that represent a denial of a victim stereotype in a GBV case (Column 1) and a non-GBV case (Column 2). The specification is the same as in our main Table 1 but adds a control for the officer's total number of key presses. The sample is smaller due to technical issues on some of the intervention days (47 (20%) females and 186 (80%) males participated). Refer also to the table notes of Table 1.

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

4.2 Robustness Checks

We find our primary estimates to be robust to an array of sensitivity checks. Overall, the findings demonstrate that the main results are robust to alternative specifications and estimations, indicating that the confrontation does indeed lead to female officers responding less stereotypically in handling a GBV case and male officers responding in the opposite direction. We start with controlling for all baseline characteristics that we found to be statistically significant in our balance checks (Table A1) and find that our main results are qualitatively very similar but less precisely estimated (Table A6). Next, Table A5 presents evidence that our main results are robust to excluding all fixed effects and additional controls and implementing heteroskedasticity-robust standard errors (not clustered). The estimates are very comparable in size and precision. Further, we control for multiple hypotheses testing and compute bootstrapped p-values for our main outcomes (Table 1, 2, and A3). Our findings are statistically

significant at conventional levels. We implement a procedure that asymptotically controls the family-wise error rate and is asymptotically balanced (Steinmayer, 2020). It is based on the approach of List et al. (2019) but modified to be used in a multivariate regression setting.

5 Conclusion

This paper has two central objectives. The first is to investigate how officers change their behavior in handling GBV cases after a prejudice confrontation. The second is to examine how treated and control officers differ in the degree of bias based on victim appearance. To achieve these objectives, we collaborated with the Madhya Pradesh police in India and conducted a confrontation-based lab-in-the-field experiment.

We find that the confrontation leads female officers to respond less stereotypically in handling a GBV case while men tend to have a counter-reaction. Females significantly prioritize the victim's statement more than the control group and give less priority to the offender's statement. Confronted females also have a higher probability of pursuing a GBV complaint (instead of dropping it) than the control group. For male officers, there is no statistically significant effect for dropping a complaint. However, the confrontation seems to come to a backlash effect on other outcomes: male officers attach lesser priority to the victim's statement and more to the offender's statement. We find no effect on outcomes, that represent officers' general disposition toward GBV, such as the belief in the truthfulness of GBV complaints and registering a GBV complaint in the first place. A potential explanation for our findings is that females predominantly show a mild bias in handling a GBV case before treatment, while more than half of the males are strongly biased. Considering that policing is highly male-dominated, the average female officer perceives a bias in her work environment that is stronger than her own. The confrontation treatment thus de-biases female officers and makes them adjust their beliefs in a professional context. In the computerized negative stimuli reaction task, we find significantly greater use of GBV stereotypes for men after a confrontation, while there is no effect for female officers. This finding is consistent with our results concerning the GBV case handling. While men seem to push back after a confrontation, women do not change their inner beliefs.

Taken together, the intervention encouraged women to de-bias their actions, which means less stereotyping when working on GBV cases. Men, on the other hand, are less receptive to the

feedback. Therefore, the task of future research is to carefully design interventions accounting fur underlying extent of biases and possible gendered-effects. On a cautionary note, our findings do not necessarily allow us to draw general conclusions about the effectiveness of confrontation in other countries as response to such interventions can be critically contingent on historical, social, and cultural factors.

References

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–59.

Alesina, A., Carlana, M., La Ferrara, E., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. *National Bureau of Economic Research*.

Amaral, S., Dahl, G. B., Hener, T., Kaiser, V., and Rainer, H. (2023). Deterrence or backlash? arrests and the dynamics of domestic violence. Technical report, National Bureau of Economic Research.

Anderberg, D., Rainer, H., Wadsworth, J., and Wilson, T. (2016). Unemployment and domestic violence: Theory and evidence. *Economic Journal*, 126(597):1947–1979.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., and Sulaiman, M. (2020). Women's empowerment in action: evidence from a randomized control trial in africa. *American Economic Journal: Applied Economics*, 12(1):210–59.

Belknap, J. (2010). Rape: Too hard to report and too easy to discredit victims. *Violence Against Women*, 16(12):1335–1344.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The quarterly journal of economics*, 126(1):103–143.

Chaney, K. E. and Sanchez, D. T. (2018). The endurance of interpersonal confrontations as a prejudice reduction strategy. *Personality and Social Psychology Bulletin*, 44(3):418–429.

Chaney, K. E., Sanchez, D. T., Alt, N. P., and Shih, M. J. (2021). The breadth of confrontations as a prejudice reduction strategy. *Social Psychological and Personality Science*, 12(3):314–322.

Chaney, K. E., Young, D. M., and Sanchez, D. T. (2015). Confrontationâs health outcomes and promotion of egalitarianism (c-hope) framework. *Translational Issues in Psychological Science*, 1(4):363.

Czopp, A. M., Monteith, M. J., and Mark, A. Y. (2006). Standing up for a change: Reducing bias through interpersonal confrontation. *Journal of personality and social psychology*, 90(5):784.

Degener, T. and Koster-Dreese, Y. (1995). Declaration on the elimination of violence against women: by general assembly resolution 48/104 of 20 december 1993. In *Human Rights and Disabled Persons*, pages 416–422. Brill Nijhoff.

García-Moreno, C., Zimmerman, C., Morris-Gehring, A., Heise, L., Amin, A., Abrahams, N., Montoya, O., Bhate-Deosthali, P., Kilonzo, N., and Watts, C. (2015). Addressing violence against women: a call to action. *The Lancet*, 385(9978):1685–1695.

Green, D. P., Wilke, A. M., and Cooper, J. (2020). Countering violence against women by encouraging disclosure: A mass media experiment in rural uganda. *Comparative Political Studies*, 53(14):2283–2320.

Guarnieri, E. and Rainer, H. (2021). Colonialism and female empowerment: A two-sided legacy. *Journal of Development Economics*, 151:102666.

Heß, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal*, 17(3):630–651.

Hidrobo, M., Peterman, A., and Heise, L. (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: evidence from a randomized experiment in northern ecuador. *American Economic Journal: Applied Economics*, 8(3):284–303.

Howell, J. L., Redford, L., Pogge, G., and Ratliff, K. A. (2017). Defensive responding to iat feedback. *Social Cognition*, 35(5):520–562.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22:773–793.

Mallett, R. K. and Wagner, D. E. (2011). The unexpectedly positive consequences of confronting sexism. *Journal of Experimental Social Psychology*, 47(1):215–220.

Palermo, T., Bleck, J., and Peterman, A. (2014). Tip of the iceberg: reporting and gender-based violence in developing countries. *American journal of epidemiology*, 179(5):602–612.

Prakash, N., Micaela, S. M., Girija, B., and Sofia, A. (2024). Debiasing Law Enforcement Officers: Evidence from an Expressive Arts Intervention in India. .

Sardinha, L., Maheu-Giroux, M., Stöckl, H., Meyer, S. R., and García-Moreno, C. (2022). Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *The Lancet*, 399(10327):803–813.

Steinmayer, A. (2020). MHTREG: Stata module for multiple hypothesis correction. *Statistical Software Components S458853, Boston College Department of Economics*.

Stevenson, B. and Wolfers, J. (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *The Quarterly Journal of Economics*, 121(1):267–288.

Tur-Prats, A. (2019). Family types and intimate partner violence: A historical perspective. *Review of Economics and Statistics*, 101(5):878–891.

Appendix A: Additional Tables and Figures

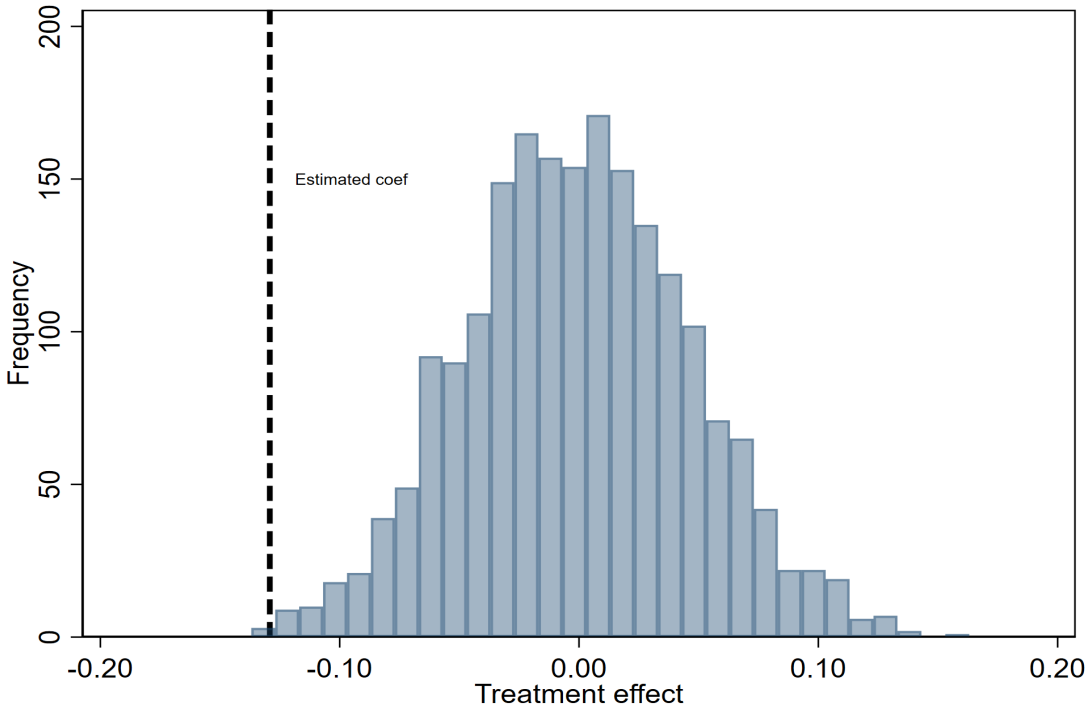


Figure A1: Randomization test

Notes: This Figure shows the result of a randomization test (ritest: Hess, Simon, 'Randomization inference with Stata: A guide and software' Stata Journal 17(3) pp. 630-651.). If the confrontation treatment had not affected any officer's case-handling at all, then the observed outcome for our main variable (i.e. prioritizing the victim's statement) would have occurred with a probability of below 0.2%, as measured by the p-value (0.002).

Table A1: Balancing tests

Variable	(1) Mean treated men	(2) Mean control men	(3) Mean treated women	(4) Mean control women	(5) Diff treated vs control men	(6) Diff treated vs control women	(7) Diff treated men vs treated women
Officer age	45.152	44.543	31.440	31.909	-0.540	-0.554	-9.383***
	(9.687)	(9.414)	(7.030)	(8.431)	(1.321)	(2.161)	(2.626)
Officer posting in Bhopal (=1)	0.420	0.394	0.320	0.273	0.119	0.032	-0.119
	(0.496)	(0.491)	(0.476)	(0.452)	(0.085)	(0.161)	(0.254)
GBV case 1: prio on victim (=1)	0.109	0.324	0.222	0.368	-0.214***	0.144	0.105
	(0.315)	(0.471)	(0.441)	(0.496)	(0.072)	(0.294)	(0.127)
GBV case 1: response time [seconds]	222.184	215.766	201.661	184.330	-7.182	29.721	1.759
	(90.360)	(95.273)	(98.161)	(128.191)	(23.536)	(27.710)	(45.451)
GBV case 2: prio on victim (=1)	0.351	0.464	0.438	0.500	-0.053	-0.302*	-0.022
	(0.481)	(0.503)	(0.512)	(0.519)	(0.118)	(0.145)	(0.208)
GBV case 2: response time [seconds]	231.783	201.703	171.268	172.088	32.871	-0.590	-79.033
	(117.034)	(105.139)	(60.482)	(59.732)	(35.745)	(22.505)	(45.244)
Non-GBV case q1: investigation of offender (=1)	0.643	0.661	0.640	0.788	0.036	-0.125	-0.072
	(0.481)	(0.475)	(0.490)	(0.415)	(0.070)	(0.195)	(0.102)
Non-GBV case q1: response time [seconds]	335.614	283.763	259.262	330.310	35.977*	-48.319	-91.789*
	(182.296)	(136.544)	(123.890)	(237.295)	(20.416)	(56.253)	(50.696)
Non-GBV case q2: prio on victim (=1)	0.938	0.921	0.920	0.970	0.034	-0.057	-0.077
	(0.243)	(0.270)	(0.277)	(0.174)	(0.049)	(0.066)	(0.092)
Non-GBV case q2: response time [seconds]	71.293	65.774	53.542	56.433	5.369	-7.202	-19.587**
	(44.921)	(40.579)	(23.140)	(40.938)	(4.113)	(6.626)	(7.780)
Senior officer tone during callout	2.080	2.142	2.240	2.242	0.004	0.000	-0.022
	(0.304)	(0.431)	(0.523)	(0.561)	(0.004)	(0.000)	(0.024)
Observations	112	127	25	33	239	58	137

Notes: This Table compares the treatment and control group in terms of their officer characteristics and baseline outcomes-both within gender (Columns 5 and 6) and across gender dimensions (Column 7). The displayed differences are regression adjusted using baseline session and senior officer as controls and cluster at baseline session level. The values in Columns (1)-(4) are simple means. 'Officer age' is the participant's age in years, 'Officer posting in Bhopal' is a dummy that takes the value 1 if the officers serves in the capital Bhopal, 'GBV case 1 or 2: prio on victim' is a dummy that takes the value 1 if participants prioritized the victim's statement (ranked as most important) in the GBV case 1 or 2, 'GBV case 1 or 2: response time' is the time it took the participants to answer to the survey question on prioritizing statements measured in seconds, 'Non-GBV case q1: investigation of offender' is a dummy that takes the value 1 if the officer ranks the investigation the offender first (instead of the victim or other people involved) in the non-GBV case, 'Non-GBV case q1: response time' measures the response time for ranking the investigation parts in the non-GBV case, 'Non-GBV case q2: prio on victim' is a dummy that takes the value 1 if the victim's statement is ranked as most important, 'Non-GBV case q2: response time' measures the officers' time to respond to the Non-GBV case q2 question, and 'Senior officer tone during callout' is a categorial variable that was coded by the research team during the feedback session with the senior officer and indicates whether the tone of the senior officer was 1:rebuking 2:neutral 3:explanatory 4:reading from script.
*p<.10, **p<.05, ***p<.01

Table A2: The effects of confrontation on the recognition of crime

	Considered as crime
	(1)
Panel A: Without gender heterogeneity	
GBV treatment	0.014
	(0.028)
Controls	yes
Control Group Mean	0.949
Rsq	0.050
N	293
Panel B: With gender heterogeneity	
GBV treatment	-0.002
	(0.036)
treatment x female	0.081
	(0.054)
female officer	-0.011
	(0.046)
Controls	yes
Control Group Mean	0.949
Rsq	0.058
N	293

Notes: This Table reports estimates of the effect of the intervention on whether the officer recognizes that a crime has committed. 'GBV treatment' is an indicator for being confronted with one's own bias in the GBV case handling. 'treatment x female' is an interaction term of this indicator and the gender dummy (=1 if female). The specification includes senior officer and session fixed effects, and officer characteristics as controls (officer's age and a dummy for posting is in the capital Bhopal). Standard errors are clustered at the session level and are reported in parentheses. The sample is reduced to 293 officers, as four of them did not answer the corresponding survey question.
*p<.10, **p<.05, ***p<.01

Table A3: The effects of confrontation on the handling of GBV cases 2

	Truth of rape complaints		Register complaint	
	(1)	(2)	(3)	(4)
Panel A: Without gender heterogeneity				
GBV treatment	-0.654	-0.686	0.034	0.038
	(0.454)	(0.448)	(0.032)	(0.029)
	[0.132]	[0.132]	[0.305]	[0.305]
Controls	no	yes	no	yes
Control Group Mean	5.325	5.325	0.931	0.931
Rsq	0.058	0.080	0.066	0.092
N	297	297	297	297
Panel B: With gender heterogeneity				
GBV treatment	-0.677	-0.706	0.044	0.048
	(0.403)	(0.411)	(0.039)	(0.036)
	[0.095]	[0.095]	[0.285]	[0.285]
treatment x female	0.131	0.112	-0.050	-0.050
	(0.813)	(0.870)	(0.040)	(0.038)
	[0.700]	[0.700]	[0.285]	[0.285]
female officer	0.877**	0.517	0.067**	0.036
	(0.364)	(0.501)	(0.023)	(0.026)
Controls	no	yes	no	yes
Control Group Mean	5.325	5.325	0.931	0.931
Rsq	0.071	0.084	0.074	0.094
N	297	297	297	297

Notes: This Table reports estimates of the effect of the intervention on the belief in the truthfulness of rape complaints (Columns 1 and 2) and the likelihood of registering a GBV complaint (dummy=1 if the complaint is registered) against the accused (Columns 3 and 4). The truthfulness of rape complaints is measured as the number of complaints (out of 10) that are considered false. 'GBV treatment' is an indicator for being confronted with one's own bias in the GBV case handling. 'treatment x female' is an interaction term of this indicator and the gender dummy (=1 if female). The specification in Columns (1) and (3) includes senior officer and session fixed effects. In Columns (2) and (4) we add officer characteristics as controls (officer's age and a dummy for posting is in the capital Bhopal). Standard errors are clustered at the session level and are reported in parentheses. Bootstrapped p values of multiple hypothesis tests are shown in square brackets. We use the Stata command mhtreg for multiple hypotheses testing that provides a procedure that asymptotically controls family-wise error rate and is asymptotically balanced (Steinmayer, 2020). It is based on List et al. (2019) but modified to be used in a multivariate regression setting.
*p<.10, **p<.05, ***p<.01

Table A4: The effect of confrontation on the handling of GBV cases: a falsification test

	Prio on victim's statement
	(1)
GBV treatment	-0.085
	(0.063)
treatment x female	0.287***
	(0.082)
female officer	-0.154**
	(0.063)
Controls	yes
Control Group Mean	0.844
Rsq	0.093
N	239

Notes: This Table reports estimates of the effect of the intervention on the likelihood of prioritizing the victim's statement. The outcome takes the value 1 if the priority is given to the victim's statement. Male officers who showed a strong bias at baseline are removed from the sample. 'GBV treatment' is an indicator for being confronted with one's own bias in the GBV case handling. 'treatment x female' is an interaction term of this indicator and the gender dummy (=1 if female). The specification includes senior officer and session fixed effects and officer characteristics as controls (officer's age and a dummy for posting is in the capital Bhopal). Standard errors are clustered at the session level and are reported in parentheses.
*p<.10, **p<.05, ***p<.01

Table A5: Robustness: The effect of confrontation on prioritizing victim statement

	Drop complaint	Prio on victim's statement	Prio on offender's statement
	(1)	(2)	(3)
GBV treatment	0.047	-0.135**	0.088**
	(0.041)	(0.053)	(0.039)
	[0.249]	[0.012]	[0.027]
treatment x female	-0.108*	0.347***	-0.239***
	(0.059)	(0.089)	(0.074)
	[0.072]	[0.000]	[0.001]
female officer	-0.026	-0.070	0.096
	(0.049)	(0.078)	(0.066)
Controls	no	no	no
Control Group Mean	0.081	0.844	0.075
Rsq	0.017	0.045	0.031
N	297	297	297

Notes: This Table reports estimates of the effect of the intervention on the likelihood of dropping a GBV complaint (Column 1), the likelihood of prioritizing the victim's statement (Column 2) and prioritizing the offender's statement (Column 3). All three outcomes are dummies and take the value 1 if the complaint is not pursued, the priority is given to the victim's statement, and the priority is given to the offender's statement, respectively. 'GBV treatment' is an indicator for being confronted with one's own bias in the GBV case handling. 'treatment x female' is an interaction term of this indicator and the gender dummy (=1 if female). The specification does not include additional controls and robust standard errors are implemented. P-values obtained by bootstrapping with replacement with 999 replications are reported in square brackets.
*p<.10, **p<.05, ***p<.01

Table A7: The effects of confrontation on guilt and self-reflection

	Guilty	Angry	Disappointed	Regretful	Ashamed	Annoyed
	(1)	(2)	(3)	(4)	(5)	(6)
GBV treatment	-0.353	0.197	-0.133	-0.569**	0.164	-0.155
	(0.259)	(0.213)	(0.159)	(0.232)	(0.186)	(0.217)
treatment x female	-0.133	-0.604	0.081	0.057	-0.155	-0.695
	(0.610)	(0.726)	(0.355)	(0.512)	(0.329)	(0.401)
female officer	-0.073	0.611	-0.241	-0.044	0.012	0.200
	(0.398)	(0.369)	(0.365)	(0.443)	(0.196)	(0.403)
Controls	yes	yes	yes	yes	yes	yes
Control Group Mean	0.214	2.225	1.975	2.819	1.525	2.206
Rsq	0.077	0.131	0.068	0.106	0.061	0.107
N	297	297	297	297	297	297

Notes: This Table shows our main specification (Panel B of Table 1) and the outcomes feeling 'guilty' (Column 1), 'angry at myself' (Column 3), 'regretful' (Column 4), 'ashamed' (Column 5), and 'annoyed' (Column 6). 'Guilty' is coded up as an index, specified by Chaney et al. (2021); the remaining variables are categorical variables (scale: 1 Not at all, 2 Doesn't apply to a great extent, 3 Applies to some extent, 4 Neither nor, 5 Somewhat applies, 6 Applies to a great extent, 7 Very much). Refer also to the table notes of our main Table 1.
*p<.10, **p<.05, ***p<.01

Table A6: Robustness: The effect of confrontation on prioritizing victim statement-including baseline controls

	Drop complaint	Prio on victim's statement	Prio on offender's statement
	(1)	(2)	(3)
Panel A: Including Baseline non-GBV case controls			
GBV treatment	0.036	-0.133**	0.097*
	(0.033)	(0.055)	(0.048)
treatment x female	-0.119**	0.378***	-0.258***
	(0.052)	(0.082)	(0.062)
female officer	0.021	-0.151**	0.131**
	(0.050)	(0.070)	(0.048)
Control Group Mean	0.081	0.844	0.075
Rsq	0.079	0.089	0.104
N	297	297	297
Panel B: Including Baseline non-GBV case + GBV case 1 controls			
GBV treatment	0.016	-0.105*	0.089*
	(0.041)	(0.057)	(0.047)
treatment x female	-0.112*	0.295***	-0.183**
	(0.063)	(0.079)	(0.075)
female officer	0.001	-0.073	0.072
	(0.055)	(0.072)	(0.066)
Control Group Mean	0.081	0.844	0.075
Rsq	0.143	0.121	0.177
N	154	154	154
Panel C: Including Baseline non-GBV case + GBV case 2 controls			
GBV treatment	0.091*	-0.178	0.087
	(0.049)	(0.105)	(0.081)
treatment x female	-0.100	0.436**	-0.336**
	(0.085)	(0.151)	(0.116)
female officer	-0.012	-0.230*	0.242**
	(0.058)	(0.118)	(0.100)
Control Group Mean	0.081	0.844	0.075
Rsq	0.177	0.180	0.184
N	143	143	143

Notes: This Table presents estimates for the same specification as in Panel B of Table 1, but includes Baseline outcomes that appeared statistically significant in our balance checks (Table A1). In Panel A, we control for the following variables: officer's age, posting, and response time for ranking parts of the investigation in the non-GBV case. In Panel B, we control for the same set of variables but add the variable 'GBV case 1: prio on victim' (whether the officer prioritizes the victim's statement in the GBV case). The sample is smaller once this variable is added, as only half of the participants handled this case at baseline (every participant handles one of two GBV cases). In Panel C, the Baseline controls on the non-GBV case from Panel A and the variable 'GBV case 2: prio on victim' (whether the officer prioritizes the victim's statement in the GBV case) are included. Refer to the table notes of Table 1 and A1.
*p<.10, **p<.05, ***p<.01

Table A8: The effects of confrontation on victim blaming and perceived bias

	Victim blaming husband beats	Victim blaming harassment	Victim blaming rape	Bias of friends	Bias of family	Bias of partner
	(1)	(2)	(3)	(4)	(5)	(6)
GBV treatment	-0.407	-0.384	0.104	-0.168	-0.087	-0.015
	(0.265)	(0.339)	(0.273)	(0.167)	(0.138)	(0.128)
treatment x female	0.508	0.073	-0.489	0.307	0.294	0.225
	(0.884)	(0.733)	(0.729)	(0.372)	(0.272)	(0.290)
female officer	0.532	0.650	0.441	-0.282	-0.226	0.045
	(0.707)	(0.576)	(0.633)	(0.194)	(0.190)	(0.178)
Controls	yes	yes	yes	yes	yes	yes
Control Group Mean	3.675	2.587	5.544	2.094	2.094	1.744
Rsq	0.107	0.102	0.105	0.124	0.093	0.090
N	297	297	297	297	297	297

Notes: This Table shows our main specification (Panel B of Table 1) and the outcomes 'victim blaming when husband beats up his wife' (Column 1), 'victim blaming when a woman is harassed' (Column 3), 'victim blaming when a woman is raped' (Column 4), where the officer is asked to indicate in how many (out of 10) cases she thinks it is the women's fault, respectively. Columns (4)-(6) show the officers' perceived bias among friends, family, and their partner, respectively, where participants were asked to what extent their friends, family, and partner think that gender based violence is the women's fault (scale: 1 0%, 2 About 25%, 3 About 50%, 4 About 75%, 5 100%). Refer also to the table notes of our main Table 1.
*p<.10, **p<.05, ***p<.01