The Effect of Manager Gender and Performance Feedback: Experimental Evidence from India

MARTIN ABEL
Bowdoin College

DANIEL BUCHMAN
Middlebury College

I. Introduction

As in many other parts of the world, women in India are closing the education gap: 53% of undergraduate degrees and 42% of PhD degrees in India were awarded to women in 2019 (Catalyst 2019). Yet, we observe the gender gap is even larger for leadership positions in the private sector, where women occupy 16% of midlevel management positions and 4% of senior-level positions (Woetzel et al. 2018). These massive gender differences in (senior) leadership positions despite a convergence of educational attainment suggest an inefficient allocation of talent. Aside from gender-equity concerns, this also hampers economic growth (Hsieh et al. 2019). Negative economic consequences may be particularly severe in developing countries such as India where firms often lack access to effective managers (Bloom, Sadun, and Van Reenan 2012; Bloom et al. 2013; McKenzie and Woodruff 2015).

Popular press narratives, personal anecdotes, and a large body of research documents accounts of gender discrimination across India's labor market (Budhwar, Saini, and Bhatnagar 2005; Zimmermann 2012; Batra and Reio Jr. 2016; Das 2000). There are few studies, however, that systematically quantify the degree of gender discrimination, and those that do tend to focus on

An earlier version of this paper that used a smaller sample size was previously circulated. This paper greatly benefited from discussions with and comments by Willa Brown, Tanya Byker, Jeff Carpenter, Siddharth George, Erik Gong, Caitlin Myers, Oyebola Okunogbe, Akhil Rao, Emcet Tas, and Maya Woser. This study was conducted in partnership with the World Bank's South Asia Gender Innovation Lab. The experiment was registered in the American Economic Association RCT Registry (no. AEARCTR-0003810), and institutional review board approval was obtained from Middlebury College. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Declarations of interest: none. All errors and omissions are our own. Contact the corresponding author, Martin Abel, at m.abel@bowdoin.edu.

Electronically published September 4, 2024

Economic Development and Cultural Change, volume 73, number 1, October 2024.

https://doi.org/10.1086/727513

job candidates or entry-level workers (Batra and Reio Jr. 2016). We know less about the prevalence and nature of discrimination toward women as they begin to occupy leadership positions. As the share of women in such positions increases—both in India and around the world—understanding potential barriers female managers face in their career advancement becomes all the more important. This paper tests whether gender discrimination by subordinates, through effort provision and reactions to criticism, presents a barrier for female managers in India.

A critical component of discrimination by subordinates is the way in which workers react to feedback. An effective manager must be able to deliver feedback that elicits positive attitude and behavior change. Yet, the evidence on the effectiveness of feedback is mixed, suggesting that such effects are highly context-specific. This has two implications: first, it highlights the need to understand further the underlying mechanisms that may explain these conflicting results, and second, it raises questions about the generalizability of existing studies. Underscoring this point is the fact that most evidence on feedback effects comes from Western contexts.

This study aims to fill these gaps and provide some of the first evidence on the effect of feedback and on the effect of manager gender in a non-Western context. We recruit a sample of 2,280 Indian workers online through Cloud Research (which recruits from Amazon's Mechanical Turk) for a transcription task and randomize the gender of the workers' (fictitious) manager as well as whether they receive feedback on their performance. This design replicates Abel (2024), which uses a sample of online workers in the United States and will serve as a benchmark for our results. The gig economy[1] is an important and growing component of India's economy. Currently, there are almost 500,000 working in data entry in India, the type of work studied in this paper. This number is projected to increase to more than 1 million Indian workers (BCG 2021).

Our study contributes four main results: First, in the absence of feedback, we find suggestive evidence that workers provide higher levels of effort working under a manager of the opposite sex. This contrasts with results from the United States where the interaction of worker and manager gender did not matter. Second, feedback receipt leads to an increase in effort provision for both mandatory and voluntary tasks. This is driven by the positive effects of negative feedback (i.e., criticism) on effort provision, which we observe despite a slight deterioration in worker attitudes when criticized. This contrasts with findings

[1] The International Labor Organization categorizes two types of gig work: microtasks that can be performed online (crowdwork), and services enabled through online platforms (BCG 2021).

from the United States showing that criticism has a large negative effect on attitudes and small (insignificant) negative effect on effort. These differences across context are significant at the 1% level. One explanation is that Indian workers perceive criticism to be more accurate and appropriate. By contrast, in both countries, praise does not affect attitudes or effort provision, which could be explained by workers' expectations of praise. Indeed, we find that workers in both contexts assess positive performance feedback to be highly accurate and appropriate.

Third, we find the effect of feedback does not vary by manager gender. Workers respond similarly to feedback from female and male managers for both praise and criticism. These results are in stark difference (significant at the 1% level) to the US sample, where attitudes deteriorate twice as much when workers are criticized by women. Fourth, because the above results may be surprising given the frequency of accounts of gender discrimination in India, we provide evidence on mechanisms that may explain the apparent absence of gender discrimination in our setting. We find no or limited evidence for the importance of attention discrimination (Bartos et al. 2016), implicit gender bias (Greenwald, Nosek, and Banaji 2003), and gendered expectations (Carli 2001; Manne 2017) in our setting. Workers spend the same amount of time reading and thinking about the feedback of female managers. In a test of the degree to which workers implicitly associate men with careers and women with family, we observe a modest level of implicit gender bias. Importantly, similar to results from the US study, implicit bias does not negatively affect workers' performance or attitude when paired with a female manager. We also do not find that workers hold explicit biases about the qualification of female managers. Finally, workers are much less likely than their US counterparts to associate certain management styles with a specific gender.

This paper relates to several strands of literature. We add to an established literature on the effects of feedback. Some papers find negative effects on productivity (Eriksson, Poulsen, and Villeval 2009; Kuhnen and Tymula 2012) while others find positive ones (Azmat and Iriberri 2010; Blanes i Vidal and Nossol 2011; Charness, Masclet, and Villeval 2014; Bandiera, Larcinese, and Rasul 2015). We contribute novel evidence from the growing gig economy and shed light on some of the underlying mechanisms, further explaining why there may be conflicting evidence across different settings.

Many cultural differences affect workplace relationships and communication, including community orientation, comfort with ambiguity, hierarchy, and so on (Amba-Rao et al. 2000; Sully De Luque and Sommer 2000; Varma, Pichler, and Srinivas 2005). It is plausible, therefore, that these differences affect how people deliver and respond to feedback. One important example is the

prevalence of gendered expectations. There is ample and very robust evidence showing that people in the United States and other Western countries strongly associate men with criticism and women with praise (Carli 2001). Importantly, we see a different pattern of associations among gig-economy workers in India. For example, strict expectations are equally associated with female and male managers. Receiving critical feedback from a woman therefore does not create a dissonance with expectations. The lack of defensiveness is also reflected by the fact that workers assess (negative) feedback from female and male managers as equally accurate and appropriate. This may be one of the explanations for why we observe less gender discrimination toward critical female managers than in the US sample (Abel 2024).

Our study also adds to a nascent literature investigating "discrimination from below" (Ayalew, Manian, and Sheth 2021; Chakraborty and Serra 2024). Existing research mainly employs lab experiments as well as evidence from survey data to explore subordinates' reactions to female managers.[2] This literature has yet to arrive at a firm consensus. Some research finds that workers, particularly men, are less likely to trust the advice or follow the instructions of female managers, but other studies suggest these effects dissipate at the most senior levels (Budhwar, Saini, and Bhatnagar 2005; Grossman et al. 2016; Asad, Banerjee, and Bhattacharya 2020; Macchiavello et al. 2020; Ayalew, Manian, and Sheth 2021).

One important aspect of our study design is that it looks at short-term supervisor-subordinate relationships without face-to-face interactions, which may limit the generalizability of results to more traditional office environments.[3] However, the gig economy and remote-work agreements have become much more prominent in recent years and are expected to increase in frequency in the wake of the COVID-19 pandemic. In addition, research suggests that the type of text-based feedback we provide can be as salient as its verbal counterpart (Bracken, Jeffres, and Neuendorf 2004). Firms increasingly gain access to large amounts of data and technology that allow them to provide real-time, automated feedback (Cecchi-Dimeglio 2017). Our results suggest that in the context of remote gig-economy workers in India, this type of feedback is effective in increasing worker effort, even for voluntary tasks.

The rest of the paper is structured as follows: Section II provides background information on gender discrimination in India. Section III describes the study design, and section IV presents the main results. Section V discusses the potential mechanisms, and section VI concludes with a discussion.

[2] A related literature shows biases against female instructors in student evaluations (Boring 2016; Buser, Hayter, and Marshall 2019; Mengel, Sauermann, and Zölitz 2019).

[3] Indeed, a recent study by Macchiavello et al. (2020) finds that in Bangladesh's garment sector, female managers are initially less effective because workers have negative beliefs about their ability.

II. Background

A. Gender Discrimination in India

Indian female labor-force participation has fallen in recent years. Some attribute this to rising household incomes, but the point remains: less than a quarter of women in India work in the formal economy. Despite scoring lower on the Human Development Index and having lower per capita incomes, Bangladesh, Bhutan, and Nepal all have significantly higher female labor-force participation rates at 36%, 59%, and 83%, respectively. In fact, India's rate is comparable with those of Pakistan and Saudi Arabia, countries where women struggle to attend school and are limited in their ability to work. This remarkably low rate is reflected in the fact that women's contribution to India's gross domestic product (GDP) is approximately 18%. Woetzel et al. (2018) suggest the biggest opportunity for advancing GDP growth in the next decade will be in bringing more women into the job market.

Unfortunately, numerous barriers still exist that prevent women from entering the workforce, discrimination being chief among them. The literature on discrimination in India indicates that, from a young age, boys are given more attention and resources in the home (Barcellos, Carvalho, and Lleras-Muney 2014; Jayachandran and Pande 2017; Dhar, Jain, and Jayachandran 2018). In the workplace, some suggest India's historically paternalistic culture has led to a consistent undervaluing of women's contributions (Batra and Reio Jr. 2016). Additionally, women face unique challenges in starting small businesses (Das 2000) and continue to receive less pay despite performing better than male counterparts in certain sectors (Ranganathan and Shivaram 2021). Some of the more recent evidence finds, however, that female professionals in India report discrimination does not limit their professional opportunities (Abbas 2015).

B. Women in Management in India

The role of women in India's managerial class, particularly in the corporate sector, is growing, albeit from a very low rate. In 2005, the number of women in managerial positions ranged between 3% and 6% of administrative positions; today, it sits around 16% for midlevel management (Budhwar, Saini, and Bhatnagar 2005; Woetzel et al. 2018). Western human resource practices and the celebration of certain "success stories" indicate a shift in cultural attitudes as well, though it is worth noting such stories may be celebrated for their exceptional nature. Most public institutions fully support gender-diversity efforts, and some private-sector employers have begun gender-diversity initiatives. A recent legal mandate helped bring board-level female representation up to 11%, though that number is still low relative to those in neighboring

countries. Despite all this, women in top leadership roles are still in the minority, accounting for only 4% of senior-level management. (Woetzel et al. 2018).

Some older evidence suggests that there remains opposition toward women in leadership among men in the workforce. For example, Gupta, Koshal, and Koshal (1998) document that men perceive women to be "less self-confident, more emotional, and less objective." Along similar lines, Budhwar, Saini, and Bhatnagar (2005) find that many female managers perceive the ego of their male subordinates as an obstacle to their effectiveness as managers. More recently, however, Abbas (2015) uses survey data to show no significant negative bias against female managers. It remains unclear whether these changes in reported attitudes reflect actual societal changes. Additionally, few studies discuss whether these shifts translate into better treatment of women in leadership roles.

C. *India's Gig Economy*

India has one of the youngest workforces in the world, and many of those tech-savvy, educated millennial workers will likely seek more flexible remote-work arrangements as they enter the labor market. Ipsos (2011) found nearly a third of Indians work from home daily and more than half are frequent remote workers, and a more recent 2016 survey revealed that one in two Indians preferred telecommuting (Basu 2016). A recent "Future of Jobs" report found that Indians are a leading source of labor in the online gig economy and have a 24% share of the global online gig workforce. The vast majority of gig workers, approximately 80%, joined within the past 5 years, and though the nature of gig work varies, approximately 31% of gig workers focus on repetitive tasks similar to those used in our study design (NobleHouse 2019).

Though research is not yet definitive, the COVID-19 pandemic is likely to accelerate many of these trends in the gig economy and remote work. While some have proposed that these types of work agreements will benefit women, as they offer more flexibility (Goldin 2014), these jobs also tend to have less regulatory oversight and may thus be particularly prone to discrimination.

III. Experimental Design

A. *Recruitment and Sample Characteristics*

We recruit study participants via Cloud Research.[4] Respondents are limited to workers over the age of 18 who reside in India. We further exclude workers with

[4] Cloud Research is an online platform that recruits participants from Mechanical Turk (MTurk). MTurk allows researchers and firms to recruit workers for simple tasks, known as human intelligence tasks (HITs). MTurk has been widely used in market research and is becoming increasingly popular for academic research. However, some researchers have voiced concerns about data quality. To address

HIT approval ratings below 95% and limited previous experience (less than 50 completed HITs) in order to improve data quality and prevent bots from completing the task.[5] The HIT description simply said "Participate in transcription task" and advertised a flat-rate payment of Rs 63 (US$0.84) for a 15-minute task.[6]

A total of 2,228 Indian workers completed the HIT between January 2020 and December 2022.[7] This sample size translates into a minimum detectable effect (MDE) size of about 0.105–0.12 SD (depending on the share of explained variation across outcomes) for the aggregate effect of feedback or manager gender. For interaction terms, the MDE is 0.15–0.17 SD, and for triple interactions it is 0.21–0.24 SD. While these MDEs are approximately 10% larger than those of the original US study, the sample provides us with sufficient statistical power to detect effect sizes on the effect of critical feedback and how it differs across manager gender reported by Abel (2024).

Table 1 shows characteristics of the full sample. As in other parts of the world, Indian gig-economy workers are not representative of the overall population. Our sample is younger and more educated, with 44% of participants having completed university. Only 27% of respondents are female, which is reflective of India's overall low female labor-force participation. Additionally, 50% of respondents are members of lower castes (mostly Other Backward Caste; OBC), and 19% are members of upper castes. While we have participants from 28 Indian states (particularly Tamil Nadu), respondents are clustered in southern India where most of India's IT industry is located.

Regarding previous work experience, 75% report having had a female supervisor in the past, which is below the share with former male supervisors (92%) but still high given the relatively small share of women in managerial positions. Levels of implicit bias are relatively low with an average score of 0.2 (compared with 0.33 in the US sample) on a scale where 0 presents no implicit bias and 1 presents strong gender bias.

B. Randomization

The study design, summarized in figure 1, closely follows the design of Abel (2024). There are two stages of randomization. First, half of participants are

---

and preempt these concerns, we use the platform Cloud Research, which prescreens MTurk participants and thus leads to more reliable data (Gupta, Rigotti, and Wilson 2021).

[5] We further divided data collection into small batches of 30–100 workers posted at different times on different days thus minimizing risks that our HIT is discussed and shared in online forums.

[6] The corresponding hourly rate of US$3.35 is slightly above the average pay for MTurk workers in India. As discussed in more detail in Abel (2024), the flat rate increases reliance on workers' intrinsic motivation. Activating intrinsic motivation is one of the key challenges of successful managers.

[7] The initial data collection from about 1,700 participants was completed in 2020. Following recommendations by referees, some additional 500 observations were collected in 2022.

TABLE 1
SAMPLE CHARACTERISTICS AND BALANCE TEST

			Manager Gender			Feedback		
	Observations (1)	Mean (2)	Female (3)	Male (4)	*p*-Value (5)	Yes (6)	No (7)	*p*-Value (8)
Age	2,205	32.1	32.14	32.06	.81	32.17	32	.59
Female	2,222	.27	.27	.27	.92	.28	.26	.36
University	2,207	.44	.43	.45	.3	.45	.43	.41
Lower caste	2,228	.5	.51	.49	.33	.5	.5	1
Upper caste	2,228	.19	.2	.18	.34	.18	.21	.1
Experience female manager	1,274	.75	.75	.76	.81	.76	.74	.38
Experience male manager	1,273	.92	.92	.93	.23	.92	.93	.75
Female manager effective	961	3.17	3.14	3.2	.26	3.2	3.12	.13
Male manager effective	1,178	3.12	3.09	3.15	.14	3.12	3.12	.99
Implicit gender bias	527	.2	.19	.22	.31	.21	.19	.59
Joint significance (*p*-value)					.60			.45

Note. Columns 1 and 2 show the sample size and sample mean for each variable. Columns 3–5 compare characteristics across the random manager-gender assignment and provide a *p*-value for a test of equal means. Columns 6–8 compare characteristics across the randomized feedback delivery and provide a *p*-value for a test of equal means.

assigned to either a fictitious female or male manager. The introduction states that "Our manager [NAME] might check in with you during the task." The gender of the manager is communicated via the title (Ms. vs. Mr.) followed by either a female or male first name.[8] First names may communicate other characteristics beyond gender (Fryer Jr. and Levitt 2004). We therefore elicited associations of different Indian names with levels of education, age, caste, and location in an out-of-sample survey on MTurk. Overall, there is little variation among common Hindu first names, and gender differences are small. We matched female and male manager names so that these characteristics are balanced by manager gender. Female manager names include Anita, Seema, Sarika, and Sharmila. Male manager names include Anil, Ashok, Dinesh, and Kamal. Table 1 suggests that randomization of manager gender was successful. None of the baseline characteristics differ between the groups at statistically significant levels, and we can reject that these characteristics are jointly significant (p-value = .60).

After recipients transcribe four receipts (discussed in more detail below), 60% are randomized into receiving performance feedback. The content of the

[8] Last names were randomized to be associated with oppressed and dominant castes. We henceforth refer to these as lower or upper castes. This randomization is used to estimate the effect of manager caste in a companion paper. We used common last names including Iyer, Kaushik, Menon Sharma, Shastri, Chamar, Kori, Pasi, and Manjhi. In all of our analyses, we control for the caste of the manager. As discussed in the companion paper, differences in manager-caste perceptions are relatively small, arguably because of regional differences in familiarity with names. This limits the extent to which we can estimate interaction effects of the caste and gender of the manager.

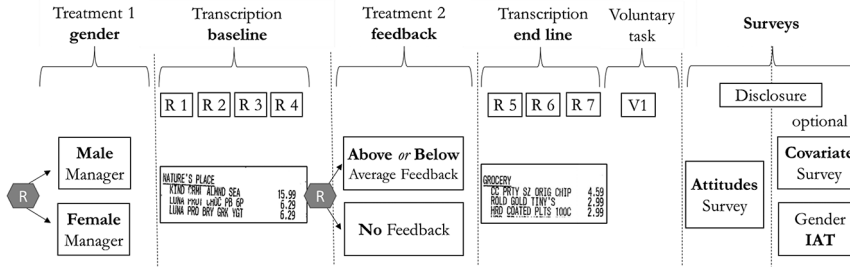

Figure 1. Design overview. IAT = implicit association test.

feedback depended on their actual performance in completing tasks associated with the initial four receipts.[9] Specifically, the text for positive/negative feedback reads:

> Hello,
>
> This is MS./MR. [NAME]. As mentioned in the task introduction, I'm overseeing your performance in transcribing the receipts.
>
> I just went over some of the receipts. Your performance has been above/below average. I was pleased with/disappointed by your effort and attention to detail.
>
> Going forward, remember that your continued commitment will improve/lack of commitment will harm the quality of our services.
>
> MS./MR. [NAME]

As discussed in more detail in Abel (2024), the text includes the standard components of feedback. Table 1 shows that characteristics are balanced between the feedback and no-feedback groups. We can reject that the variables are jointly significant (p-value = .45).

C. *Transcription Task and Outcomes*

We choose a real-effort transcription task and ask participants to copy both the item description and price from a scanned copy of an actual receipt. This task includes both language and math components and may thus be seen as more gender neutral (Niederle and Vesterlund 2007). In fact, female and male workers perform at similar levels in this task (table 2). Furthermore, these types of transcription tasks are among the most common tasks on MTurk. The introduction of the HIT states that the transcription task can "help businesses understand . . . how to spend money better so they can increase their bottom

[9] The scoring was automated so that we could provide instant feedback.

TABLE 2
EFFECT OF MANAGER GENDER ON EFFORT (BASELINE; *N* = 2,222)

	Legible		Adding		Transcription Score	
	(1)	(2)	(3)	(4)	(5)	(6)
Female manager	.004	.018	.012	.036*	.042	.245*
	(.015)	(.017)	(.018)	(.021)	(.112)	(.132)
Female worker	−.004	.023	−.009	.036	−.083	.302*
	(.017)	(.024)	(.020)	(.028)	(.124)	(.174)
Female manager × female worker		−.052		−.089**		−.746***
		(.033)		(.040)		(.247)
Sample mean	.69	.69	.57	.57	16.15	16.15
SD	.35	.35	.42	.42	2.68	2.68
p-value		.23		.12		.02

Note. The dependent variable in cols. 1 and 2 is whether the worker says the receipt is legible. The dependent variable in cols. 3 and 4 measures whether the worker is willing to add up the amounts. The dependent variable in cols. 5 and 6 captures the accuracy of transcribing receipts. All estimations are by ordinary least squares (OLS). Robust standard errors are in parentheses. *p*-Value presents results testing whether female managers have a positive effect for female workers.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

line."[10] While this statement is not false, it misleads participants about the purpose of the research, which is necessary for most discrimination studies to ensure that participants do not try to hide discriminatory attitudes and behaviors.

1. Effort Outcomes

Transcription results offer us an objective measure of workers' effort and accuracy. Similar to List and Momeni (2017), we present workers with receipts of varying readability and first ask them whether that receipt is legible (fig. 2). On average, 70% of receipts were deemed legible. This presents a measure of effort at the extensive margin. In order to also measure the quality of the work, we construct a transcription score. Each correctly transcribed item counts as 1 and each correct price as 0.5.[11] The average score is 16, which is similar to the performance of US workers. These outcomes measure effort related to mandatory tasks, as specified in the HIT introduction.

We also measure effort in voluntary tasks. For many real-world tasks, individual effort or output is hard to observe. One of the key challenges of managers is

[10] The HIT introduction does mention that "performance data will be recorded for research purposes." This type of data-collection arrangement is increasingly common in data collection from customers and also from workers.

[11] For receipts deemed illegible, we set values for effort measures (adding, transcription accuracy) to zero. Results are unchanged in magnitude and significance when we use different scoring rules, e.g., assigning 60% of possible points for skipping a receipt. This rule reflects that very poor performance is often more detrimental than not completing the task at all.

Figure 2. Receipt example.

to leverage intrinsic motivation of workers for tasks that may be hard to contractually specify (Podsakoff et al. 1990). We measure voluntary effort in two ways. For each legible receipt, we ask workers to add up all item prices. We emphasize that this is "not a required task and will not affect your payment." Overall, 57% of workers complete this task, which is similar to the percentage in our US sample (Abel 2024). One issue with this measure is that we only collect it for those who say the receipt is legible. To address potential selection issues, we introduce a second measure. After the completion of the transcription, we ask participants to complete the "completely voluntary" task of grouping receipts according to the total sum.[12] The combination of mandatory and voluntary effort allows us to test whether (negative) feedback crowds out intrinsic motivation and leads workers to reallocate effort to the incentivized outcomes similar to multitasking concerns (Holmstrom and Milgrom 1991).

2. Attitude

After the completion of the transcription but before the debrief, we collect data on workers' attitudes. Specifically, we first ask whether workers want to "sign up to do additional work for us in the next few days?" This serves as a proxy for worker retention, which is an important challenge for many firms (Hoffman, Kahn, and Li 2017). Next, we ask how much workers agree that they are (*a*) satisfied with the "experience transcribing receipts" and (*b*) convinced that "the task was important." Finally, we ask those who received feedback

[12] Specifically, the instructions read: "I need to group receipts into two categories, depending on whether the sum of all listed items is greater than 10.00. For each of the receipts below, please indicate whether the items listed add up to an amount greater than 10.00."

whether they agree that the feedback was accurate and appropriate (for the complete survey, see app. B, available online).

D. Estimation Strategy

As described, the experimental design features two stages of randomization: the gender of the manager (FemMgr) and whether workers receive feedback (Feedback). The effect of feedback and manager gender can thus be estimated through the following specification:

$$y_i = \beta_0 + \beta_1 \text{FemMgr}_i + \beta_2 \text{Feedback}_i + \beta_3 \text{FemMgr} \times \text{Feedb}_i + \epsilon_i. \quad (1)$$

Here, y_i measures outcomes y for worker i. As specified in the preanalysis plan, we collect two types of outcomes y: measures of effort (legibility, adding, transcription score, voluntary task) and attitude (interest in future work, task satisfaction, task importance). We combine the three attitude measures into a standardized index to allay multiple hypothesis testing concerns (Kling, Liebman, and Katz 2007). The term β_1 measures the effect of a female manager who does not provide feedback, β_2 is the effect of feedback from a male supervisor, and $\beta_2 + \beta_3$ represents the effect of feedback from a female supervisor.[13]

To estimate the effect of feedback content, we estimate

$$y_i = \gamma_0 + \gamma_1 \text{AboveAv}_i + \gamma_2 \text{Feedback}_i + \gamma_3 \text{AboveAv} \times \text{Feedb}_i + \epsilon_i. \quad (2)$$

Here, AboveAv is a binary variable indicating whether workers performed above average. Therefore, γ_2 measures the effect of negative feedback, and $\gamma_2 + \gamma_3$ captures the effect of positive feedback.

IV. Results

Following our preanalysis plan, we first present the effect of manager gender before workers receive any feedback (4.1), then discuss the effects of feedback (4.2) and finally look at the interaction of manager gender and feedback (4.3).

A. Manager Gender and Effort Provision

Results in table 2 test whether, before receiving feedback, manager gender has an effect on the level of effort workers exert for both mandatory and voluntary tasks. Pooling across worker gender, coefficients on the effect of being assigned a female

[13] We use names to induce variation in the perception of manager gender. While we picked names that unambiguously indicate gender, this strategy assumes that people notice the name. These coefficients should thus be interpreted as intent-to-treat estimates (akin to résumé studies). Without knowing the actual variation in perception induced by names, one cannot estimate the local average treatment effect (LATE). However, in an attempt to increase the salience of names, we reiterated this information and forced people to spend a minimum amount of time on the respective sites (following the original design).

manager are precisely estimated and very close to zero at about 0.01–0.02 SD (cols. 1, 3, and 5). However, we find suggestive evidence that the interaction of worker and manager gender affects effort provision. Male workers are slightly more likely to complete the voluntary task and have a higher score if working under a female manager. These effect sizes are modest (below 0.1 SD) and only significant at the 10% level (cols. 4 and 6). By contrast, women are less likely to add up receipts and provide less effort if assigned to a female rather than a male manager (cols. 4 and 6). This behavior contrasts with the US sample where the interaction of worker and manager gender does not affect effort provision (Abel 2024).

B. Effect of Feedback Content

1. Attitude

Table 3 shows the effects of feedback on the three attitude outcomes as well as the standardized attitude outcomes. The aggregate effect of feedback is negative but small and relatively modest in magnitude for all attitude outcomes (cols. 1, 3, 5, and 7). These aggregate effects mask important differences by feedback content. The feedback coefficients in columns 2, 4, 6, and 8 show that negative feedback leads to a drop in attitude outcomes by about 0.15 SD. This estimate is significant at the 5% level for the attitude index (col. 8). By contrast, the effect of praise is very close to zero and not statistically significant.

These results are similar across worker gender, although it is notable that the negative effect of criticism on attitude tends to be larger for male workers (table A1).

TABLE 3
EFFECT OF FEEDBACK CONTENT ON ATTITUDES (*N* = 2,201)

	Work Future		Job Satisfaction		Task Import		Attitude Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feedback	−.009	−.015	−.072*	−.120*	−.060	−.137**	−.063**	−.112**
	(.008)	(.012)	(.040)	(.062)	(.045)	(.067)	(.030)	(.047)
Above × feedback		.011		.055		.130		.072
		(.016)		(.077)		(.089)		(.059)
Above average		.010		.447***		.189***		.250***
		(.012)		(.062)		(.072)		(.046)
Sample mean	.96	.96	5.24	5.24	5.18	5.18	.00	.00
SD	.19	.19	.91	.91	1.03	1.03	.70	.70
Positive feedback = 0 (*p*-value)		.68		.16		.90		.26

Note. Work future is a binary variable whether workers are interested to work for the firm in the future. Job satisfaction and task import measure how strongly workers agree that they were satisfied with the task and that the task was important, respectively (0 = strongly disagree, 6 = strongly agree). Attitude index is a standardized index of the three outcome measures. Above average is a dummy if the worker performed above average. All estimations are by OLS. Robust standard errors are in parentheses. *p*-Values report tests of whether the effect of praise is different from zero.

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

While this difference is not statistically significant, it goes against a common narrative that women may react more negatively to criticism (Correll and Simard 2016). As discussed in more detail in section VI, it is noteworthy that the overall pattern of negative effects of criticism and no effect of praise are similar to results from the United States with the important difference that the negative effects of criticism in India are only about one-third the magnitude of those in the United States.

2. Effort

Table 4 shows the effects of feedback on effort. The aggregate effects are positive and moderate in size, ranging between 0.05 and 0.15 SD (cols. 1, 3, 5, and 7). These positive results are driven by the effects of criticism. Negative feedback increases the share of legible receipts by 5.7 percentage points (0.17 SD), increases the transcription score by 0.12 SD, and even increases the share completing the voluntary task by 4.5 percentage points (0.13 SD). This last result is particularly surprising: despite the negative effect of criticism on workers' attitude, it increases effort for both mandatory and voluntary tasks.

The effect of praise on effort is very similar to the attitude results. For both mandatory and voluntary tasks, estimates are close to zero and not statistically significant. These effects are again similar by worker gender (table A2). One explanation for the null effect of praise on both attitude and effort is that workers

TABLE 4
EFFECT OF FEEDBACK CONTENT ON EFFORT

	Legible		Adding		Transcription Score		Voluntary Task	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feedback	.029**	.057***	.019	.033	.167***	.246***	.027	.045*
	(.011)	(.019)	(.014)	(.021)	(.065)	(.088)	(.017)	(.025)
Above × feedback		−.059***		−.034		−.163		−.037
		(.022)		(.028)		(.129)		(.034)
Above average		.072***		.100***		.197		.059*
		(.023)		(.028)		(.127)		(.031)
Observations	2,201	2,201	2,201	2,201	2,201	2,201	2,064	2,064
Sample mean	.61	.61	.50	.50	12.82	12.82	.83	.83
SD	.34	.34	.39	.39	2.12	2.12	.37	.37
Positive feedback = 0 (*p*-value)		.87		.98		.38		.74

Note. The dependent variable in cols. 1 and 2 is whether the worker says the receipt is legible. The dependent variable in cols. 3 and 4 measures whether the worker is willing to add up the amounts. The dependent variable in cols. 5 and 6 captures the accuracy of transcribing receipts. Columns 7 and 8 use the score as the dependent variable and divide the sample by worker gender. All estimations are by OLS. Robust standard errors are in parentheses. *p*-Values report tests of whether the effect of praise is different from zero.

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

expect to be praised and thus do not change their attitude or effort. As discussed in more detail below, workers think that positive feedback is highly accurate and appropriate.

C. *Effect of Manager Gender*

This section looks at the effect of manager gender and feedback. The aggregate effect of feedback can be estimated through an interaction term (specification 1). To look at the interaction of manager gender and feedback content, we need to estimate models with a triple interaction of manager gender, feedback, and performance. While these results are presented in appendix A, tables A3 and A4, this section will show graphs with nonparametric results for greater transparency and ease of interpretation. This strategy effectively compares average outcomes across eight groups. One limitation is that the cell sizes are smaller (on average 275 workers), which reduces the precision of estimates.

1. Attitude

Figure 3 shows how the effect of feedback on the attitude index differs by manager gender for negative feedback (left panel) and positive feedback (right

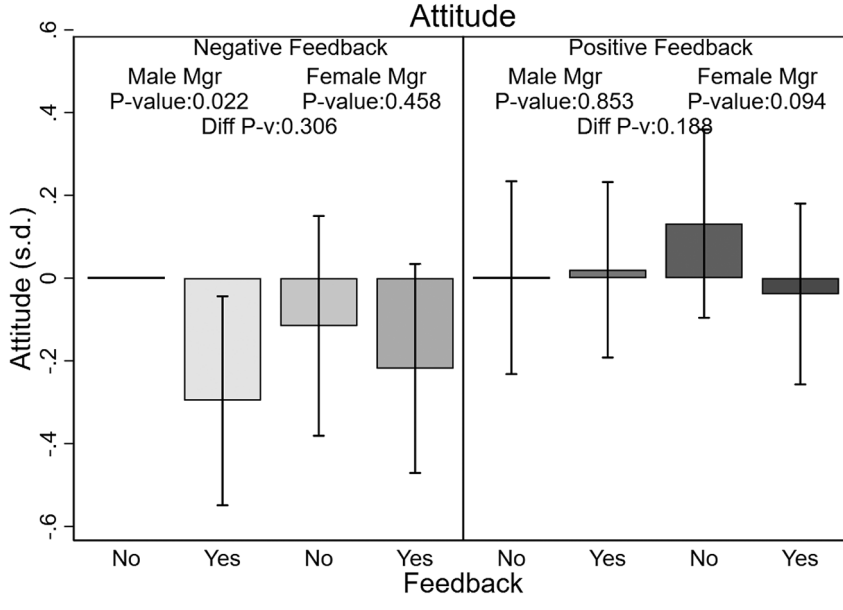

Figure 3. Attitude and feedback content. Shown is the effect of negative-feedback (left panel) and positive-feedback (right panel) participants working under male and female managers on the standardized attitude outcome. This attitude outcome is normalized so that zero presents the attitude of workers assigned to male managers and who do not receive feedback. *p*-Values are reported for a test of equal means across feedback assignment ("P-value") and to test whether the effect of feedback differs by manager gender ("Diff P-v"); 90% confidence intervals are displayed.

panel). The drop in attitudes in response to negative feedback is similar between male and female managers (p-value = .31). Similar to results from the United States, we find that for both female and male managers, positive feedback has no effect. These results are similar for the three attitude outcomes that compose the attitude index (table A3). There is also no significant difference when we look at the interaction of manager and worker gender in how feedback is received (results available on request).

2. Effort

Similar to the results on worker attitudes, there is little evidence of gender discrimination in terms of effort provision. Figure 4 shows that responses to criticism are similar between female and male managers (left panel). We also find that manager gender does not matter for the effect of criticism on voluntary tasks (table A4).

The right panel of figure 4 shows that people tend to respond more positively to praise from female managers, although these differences are not statistically

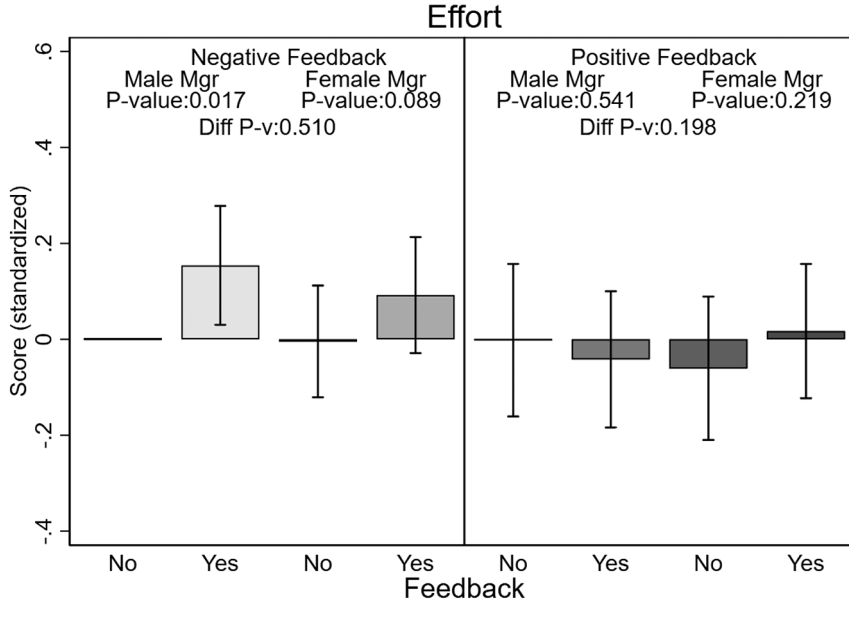

Figure 4. Effort and feedback content. Shown is the effect of negative-feedback (left panel) and positive-feedback (right panel) participants working under male and female managers on the standardized transcription score. This score is normalized so that zero presents the score of workers assigned to male managers and who do not receive feedback. *p*-Values are reported for a test of equal means across feedback assignment ("P-value") and to test whether the effect of feedback differs by manager gender ("Diff P-v"); 90% confidence intervals are displayed.

significant (p-value = .20). These results are driven by responses of male workers, but worker gender differences are also not significant (results available on request).

V. Mechanism

The previous section showed that workers do not discriminate against (critical) female managers with regard to effort or attitudes. These findings may be surprising because they differ from results in the United States (Abel 2024) and from other evidence out of India, which suggests gender discrimination poses a barrier to the career advancement of women (Zimmermann 2012; Batra and Reio Jr. 2016; Das 2000). In this section, we take advantage of the rich data we collect to better understand the underlying mechanisms that may explain why our results differ from those of other studies. As specified in the preanalysis plan, our design tests for the role of attention discrimination (sec. V.A), implicit bias (sec. V.B), and gendered expectations (sec. V.C).

A. Attention Discrimination

It has long been proposed that women are more likely to be ignored when trying to influence people (Carli 2001). This would provide a significant challenge for managers to lead and change subordinates' behavior. Evidence on the role of attention discrimination is scarce because of data-availability limitations.[14] In this study, we measure the exact time that workers spend reading and thinking about managers' feedback before continuing to the transcription task, which provides a unique opportunity to test the attention-discrimination hypothesis.

We do not find evidence for attention discrimination against women in our context. Table 5 shows that in aggregate, workers spend 0.2 seconds longer on feedback received from female managers (col. 1). This difference is statistically insignificant and small compared with the control mean of 19 seconds. The amount of attention female managers receive also does not differ by workers' gender (col. 2) or by the content of the feedback (col. 3). The same results hold when we look at the role of feedback content separately by worker gender (cols. 4 and 5). While these results are qualitatively similar to those from the US sample, it is notable that Indian workers spend on average 60% longer reading the feedback.

[14] One notable exception is Bartos et al. (2016) who find attention discrimination toward immigrants in the rental market in Germany.

TABLE 5
ATTENTION TO FEEDBACK

	(1)	(2)	(3)	(4)	(5)
Manager female	.218	.248	.163	−.008	.114
	(.687)	(.809)	(1.035)	(1.878)	(1.248)
Worker female	−.898	−.846	−.889		
	(.765)	(1.049)	(.759)		
Female manager × female worker		−.105			
		(1.531)			
Above average			3.570***	3.755**	3.449***
			(1.002)	(1.772)	(1.218)
Above × female manager			−.109	.238	−.121
			(1.373)	(2.598)	(1.631)
Observations	1,302	1,302	1,302	363	939
Sample mean	18.56	18.56	18.56	18.56	18.56
SD	12.64	12.64	12.64	12.64	12.64
Sample (workers)	Pooled	Pooled	Pooled	Female	Male

Note. The dependent variable for all columns is the time respondents spend reading the feedback from their manager (in seconds, winsorized at the 10% level). All estimations are by OLS. Robust standard errors are in parentheses.
** $p < .05$.
*** $p < .01$.

B. Implicit Gender Bias

After the end of the transcription task and survey, we offer workers a small bonus for completing a "Gender-Career" implicit bias test (IAT).[15] The IAT measures the speed with which someone associates male names with professional terms (e.g., "office," "manager," and "salary") and female names with family terms ("marriage," "home," and "children").[16] The difference in speed is transformed to a standardized score ranging between 1 and −1 where 0 indicates no bias. Slight, moderate, and strong biases correspond to scores of 0.15, 0.3, and 0.6, respectively.

The first surprising evidence is that levels of gender bias among Indian gig-economy workers are substantially lower than among MTurk workers recruited

[15] Participants were randomized to complete either the gender or caste IAT. About 70% agreed to participate, resulting in a sample of about 500 who completed the gender IAT. While selection into completing the IAT is balanced across gender, younger and more educated people are more likely to complete it. Importantly, the share completing the IAT is almost identical across our randomly assigned treatment groups (feedback and manager gender).

[16] There is an active debate to what extent implicit biases are important drivers of behavior. One side claims that IATs are poor predictors of actual behavior (Oswald et al. 2013). The other side argues that implicit biases are far more common and predictive than explicit biases (Greenwald, Banaji, and Nosek 2015; Grossman et al. 2016). For example, Reuben, Sapienza, and Zingales (2014) find that employers' IAT scores predict biased updating of expectations upon receiving performance information. Along the same lines, Glover, Pallais, and Pariente (2019) find that implicit biases among managers lower the job performance of minority workers.

from the United States. The average score in our Indian sample is 0.2 compared with 0.33 in the US sample (Abel 2024). A score of 0.2 indicates relatively low levels of implicit bias with regard to women and career.

Table 6 shows how workers' implicit gender bias is correlated with the effort and attitudes and how this relationship varies with the gender of their manager. Overall, more biased workers perform more poorly and have worse attitudes. However, this negative relationship is lower (not significant) if people are working under a female manager (cols. 1 and 5), especially when workers receive feedback (cols. 2 and 6). For effort, this relationship seems to be driven by the effects of negative feedback (cols. 3 and 4). For attitudes, the content of the feedback does not matter (cols. 7 and 8).

Overall, there is no evidence that implicit gender biases play an important role in the worker-manager relationship in our setting. If at all, more biased workers tend to react more positively if paired with a female manager. One explanation is that the gender IAT measures people's association of women and career and is thus less important once workers see women in a position of professional leadership. It is also possible that biased workers believe that

TABLE 6
IMPLICIT BIAS

	Transcription Score				Attitude Index			
	Aggregate (1)	Feedback (2)	Positive Feedback (3)	Negative Feedback (4)	Aggregate (5)	Feedback (6)	Positive Feedback (7)	Negative Feedback (8)
Gender IAT	−1.180**	−.224	−.361	−1.386	−.156	.015	−.108	.096
	(.529)	(.695)	(.546)	(1.235)	(.120)	(.202)	(.193)	(.383)
Female manager × IAT	.698	−.687	.041	−.730	.246	.075	.268	−.174
	(.679)	(.914)	(.781)	(1.667)	(.152)	(.239)	(.250)	(.433)
IAT × feedback		−1.683*	−1.098	−1.365		−.289	−.207	−.277
		(.997)	(1.079)	(1.644)		(.245)	(.253)	(.457)
Female manager × feedback × IAT		2.442*	.618	3.823*		.290	.001	.612
		(1.307)	(1.304)	(2.311)		(.306)	(.334)	(.543)
Observations	486	486	295	191	486	486	295	191
Sample mean	13.03	13.03	13.03	13.03	.01	.01	.01	.01
SD	2.327	2.327	2.327	2.327	.707	.707	.707	.707

Note. The dependent variable in cols. 1–4 is the transcription score described above. The dependent variable in cols. 5–8 is the standardized attitude index. Larger values in the IAT variable mean that respondents hold a stronger implicit gender bias. Models are fully interacted, but only coefficients of interest are reported for readability reasons. All estimations are by OLS. Robust standard errors are in parentheses.
* $p < .10$.
** $p < .05$.

female managers needed to overcome more barriers and are thus particularly qualified.

We also collected evidence on the explicit beliefs about the qualification of female managers. In an important recent study in neighboring Bangladesh, Macchiavello et al. (2020) show that garment-industry workers initially believe that female managers are of lower ability, resulting in lower performance. We collect data on workers' perception of their managers' level of education: 64.2% of workers thought their female manager attended university. This figure is very similar to the 62.4% of workers who thought their male manager attended university, suggesting that there are no gender differences in perceived manager ability.[17]

C. *Gendered Expectations and the Perception of Feedback*

People tend to react negatively when their expectations are violated.[18] This also applies to workers' reaction to feedback that contradicts expectations about their own performance (Kuhnen and Tymula 2012; Gjedrem 2018). This is important in the context of gender discrimination, as numerous studies document that people hold gendered expectations. Women are seen to be more agreeable, altruistic, warm, compliant, modest, and sympathetic than men (Eagly, Karau, and Makhijani 1995; Bertrand 2011; Blau and Kahn 2017). Abel (2024) concludes that strong gendered expectations among gig-economy workers in the United States can explain gender discrimination against critical female managers, confirming that gendered expectations have important implications for the effectiveness of female leaders (Eagly, Karau, and Makhijani 1995).

Most of the evidence on gendered expectations comes from developed countries. We elicit associations of gender and management styles through an out-of-sample survey with 141 Indian MTurk participants. Table 7 shows the share of respondents that associate certain management styles with women, men, or either. Notably, we find that for many traits, people do not hold gendered expectations. The shares associating female or male managers with strict expectations are identical. Likewise, associations for giving praise and giving feedback are very similar. For other traits, we see somewhat stronger gendered

[17] This result contrasts with that of Macchiavello et al. (2020) who find that garment-factory workers in Bangladesh falsely believe that female supervisors have less technical knowledge. One explanation is that gender stereotypes about ability are domain specific (Coffman, Flikkema, and Shurchkov 2021).

[18] It is part of a broader literature showing that people react negatively if their expectations are violated (for a review, see Burgoon 1993).

TABLE 7
MANAGEMENT STYLE EXPECTATIONS

	Observations	Female	Male	Either	M/F Ratio
Strict expectation	141	.33	.33	.34	1
Give criticism	141	.22	.35	.43	1.59
Give praise	141	.26	.29	.45	1.11
Appropriate tone	141	.36	.23	.40	.64
Communicate expectations	141	.30	.24	.46	.80
Give feedback	141	.29	.26	.45	.90

Note. Presented are results from an out-of-sample survey asking people whether they associate different management styles with female or male managers or either gender. M/F ratio reports the response ratio of male vs. female association.

expectations, although these are still much weaker than among gig workers in the United States (Abel 2024).

Weaker gendered expectations could thus be one of the reasons for why we see less gender discrimination in the Indian context. This is also reflected by how workers interpret the feedback they receive. In the United States, (male) workers were more likely to discard critical feedback from female managers as inaccurate. We do not observe the same pattern in India. The accuracy and appropriateness of feedback is almost identical across manager gender (table A5, cols. 1 and 5). In contrast to evidence from the United States, we also do not find that perception of accuracy or appropriateness of critical feedback do not vary by manager gender (table A5, cols. 2 and 6) or across worker gender (cols. 3, 4, 7, and 8).

One may be concerned that participants recruited via MTurk may change their behavior because they may suspect they are participating in research.[19] While we cannot rule out these Hawthorne effects, it is reassuring that even after being prompted, only 20% of participants expected the task to be related to research. Importantly, when asked about the exact nature of the research questions, none of the participants thought it was related to gender or the identity of the manager.

VI. Discussion

A. Comparison: India versus United States

Our study provides some of the first experimental evidence on the effect of manager gender and feedback in a non-Western context. It is, to our knowledge, also the first experimental replication on the effect of feedback across countries. Figure 5 compares our estimates to those of Abel (2024), who employs the same study design with workers in the US gig economy.

[19] MTurk is increasingly used in the United States as a platform to recruit participants for research. We are aware of fewer MTurk-based research studies in a non-Western context.

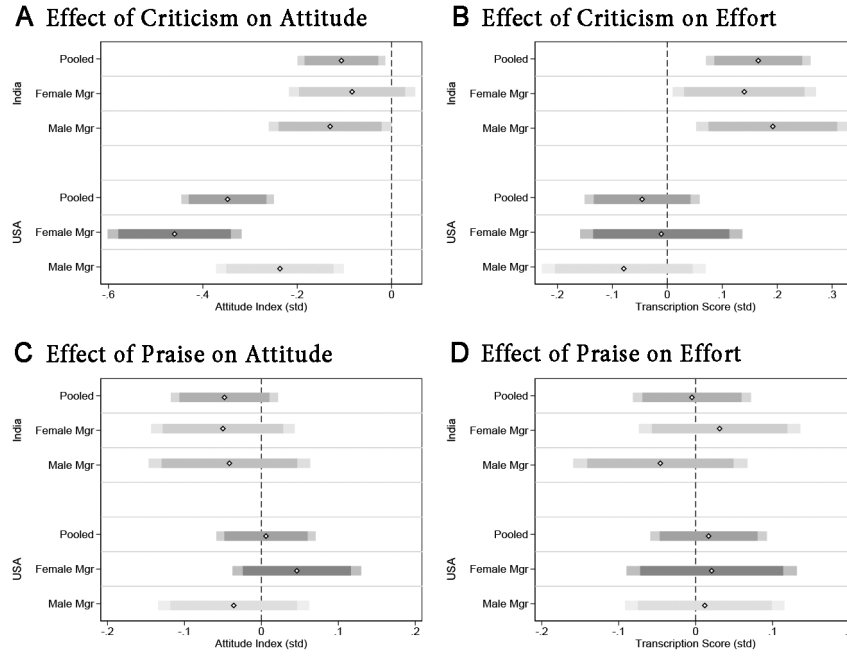

Figure 5. Effect of criticism: India versus United States. Shown are the effects of criticism (*A* and *B*) and praise (*C* and *D*) on attitude (*A* and *C*) and effort (*B* and *D*). Each point estimate reports differences in responses relative to the group of workers that did not receive feedback. Results are presented for both the pooled sample and separately by manager gender for both India and the United States. Outcomes are standardized (std) to have a mean of zero and standard deviation of one. The bars present 90% and 95% confidence intervals.

The pooled results show that praise has no effect on attitude or effort in either the United States or India (figs. 5*C* and 5*D*). By contrast, the negative effect of criticism on attitude is about three times larger in the US sample (0.45 SD) than in the Indian counterpart (0.15 SD). Moreover, criticism decreases effort among workers in the US sample (0.1 SD) but increases it among Indian counterparts (0.17 SD). These cross-country differences are sizable in magnitude and statistically significant at the 1% level.

These findings suggest that Indian workers in our sample not only react less poorly to critical feedback but also increase effort provision in response to being criticized. This is consistent with research on the effects of India's high-status-identity, high-power-distance culture on workplace dynamics. This increased sense of hierarchy can result in greater inherent expectations of top-down feedback delivery (Sully De Luque and Sommer 2000) and respect for paternalistic authority (Amba-Rao et al. 2000). Additionally, Kwantes (2009) suggests that loyalty plays a much bigger role in Indian workers' commitment to their organizations, which implies effort provision and attitudes toward supervisors are

driven by a sense of obligation rather than a subjective evaluation of their job satisfaction.

In our study, the increased acceptance of hierarchy is reflected in Indian workers' perception of feedback. Figure 6 shows the perceived accuracy of feedback as a function of workers' performance in India (fig. 6*A*) and the United States (fig. 6*B*). Specifically, we measure how much people agree with the statement that the feedback they received was accurate and record responses as 0 for "strongly disagree" through 6 for "strongly agree." Praise is perceived as highly accurate in both countries, as depicted by the high levels of perceived feedback accuracy to the right of the average threshold. In both contexts, criticism is perceived as less accurate than praise. Levels of agreement with criticism, however, vary substantially. On average, US workers "somewhat disagree" with the statement, while Indian workers on average "somewhat agree" or "mostly agree" that the criticism they received was accurate. Results for the perceived appropriateness of feedback are very similar.

Another notable cross-country difference is in the role of manager gender. Criticism has a small negative effect (0.12 SD) on attitudes in India, regardless of manager gender. By contrast, in the United States, the effect of negative feedback is twice as large if it comes from a female manager (fig. 5*A*). These differences in worker responses to critical female managers is significant at the 1% level. Consistent with this pattern, we observe that Indian workers do not assess (negative) feedback to be less accurate if it comes from a female manager, which contrasts with results from the US sample, where workers think criticism is less accurate when it comes from a female manager.

Across both countries, we do not find gender discrimination in response to praise. US workers react slightly more positively to praise from women with

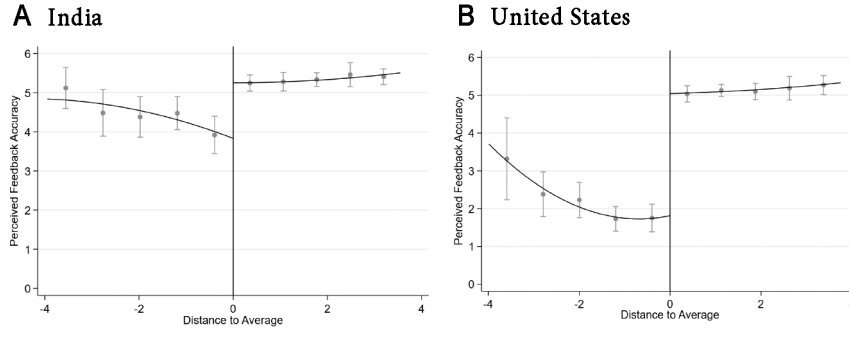

Figure 6. Perception of feedback: India versus United States. Shown is the perceived accuracy of feedback in India (*A*) and the United States (*B*). The *x*-axis measures the transcription score. Feedback to the left (right) of the average is negative (positive). The *y*-axis measures whether workers say that the feedback was accurate, ranging from 0 = strongly disagree to 6 = strongly agree.

respect to attitudes, but this difference is not significant, and there is no differential response in effort provision. By contrast, Indian workers tend to increase effort more if praised by a female manager compared with a male manager, but these differences are also not statistically significant (figs. 5*C* and 5*D*).

Our discussion of the underlying mechanisms highlights that gender discrimination and other aspects of workplace culture are highly context-specific. One's culture and belief system directly inform the gendered expectations one holds; one's upbringing and environment can change the expectations one has of a particular gender, and one's learned habits may be directly responsible for one's likelihood to pay attention to or ignore a person of a particular gender. The benefit of replicating study designs with comparable samples across countries is that it can help to identify context-specific factors that determine how workers respond to various treatments. This is particularly useful in studying culturally variable phenomena such as gender-based discrimination from below.

B. External Validity

The specific work arrangement we are investigating, remote work in the gig economy, is growing in popularity. Particularly in the wake of the COVID-19 pandemic, Indian workers, who had long expressed a growing interest in telecommuting, are likely to pursue more remote-work arrangements as well as flexible gig-based work (Ipsos 2011; Basu 2016). Few studies have explored the unique nature of the supervisor-subordinate relationship in remote-work arrangements in India or other parts of the Global South more broadly.

While studying this growing sector is important in its own right, it is also important to acknowledge that our sample is not representative of the Indian workforce. Gig-economy workers are disproportionately young, male, and university educated. To test the extent to which this selection affects the generalizability of our findings, we explore how our results vary across relevant subgroups. We find the effect of criticism and praise on attitudes is consistent across all groups, and none of the differences in attitude effects exceed 0.1 SD. There is slightly more variation in the effect of negative feedback on effort. Most notably, older respondents react more negatively to criticism from female managers. A similar pattern holds for members of upper castes, although given the relatively small sample size, this effect is estimated imprecisely.

Another concern is that our sample is predominantly drawn from India's South. Nearly 65% of workers in our sample reside in either Tamil Nadu or Kerala, reflecting that most of India's IT industry is located in the South. Traditionally, these states tend to be more gender progressive, have fewer gender disparities in education, and boast higher levels of female labor force participation and female autonomy (Rammohan and Vu 2018). Comparing results

between workers in the South and other parts of India indeed shows that there tends to be less gender discrimination in response to criticism in the South. These differences, however, are relatively small and not statistically significant.[20]

Overall, these results suggest there is relatively little variation across subgroups, which partially allays concerns over the generalizability of our results. Nevertheless, we want to reiterate that we do not make any wider claims about whether results generalize to other parts of India's economy. Indeed, many studies show that gendered expectations and gender discrimination are still prevalent in Indian society (Zimmermann 2012; Barcellos, Carvalho, and Lleras-Muney 2014; Batra and Reio Jr. 2016; Das 2000; Ranganathan and Shivaram 2021). Even if one does not accept the premise that in-person gig work and remote environments are full of the exact types of interactions studied in this paper, there are almost 500,000 Indians doing the type of data-entry work we look at, and this number is projected to increase to more than 1 million workers (BCG 2021).

There is, however, evidence that gender discrimination may be on the decline in India. Abbas (2015) finds that gender expectations in India have been converging over time, and Beaman et al. (2012) show that the increase of women serving as village leaders led to a reduction in gender-biased attitudes and an increase in support for female political leaders. By contrast, there is little evidence on the interplay of hierarchical norms and traditional gender expectations in India's private sector. Our study presents some initial evidence to suggest that deference to hierarchy may play a role in mitigating the effect of traditional gender norms on workplace interactions.

C. Concluding Remarks

This study closely replicates in India the recruitment and experimental design of a study first conducted in the United States. We find important differences in workers' behavior and attitudes, which highlight the importance of cultural factors in how workers perceive and react to performance feedback. Specifically, negative feedback has a less negative effect on attitude and more positive effect on effort provision among Indian workers. In contrast to their US counterparts, Indian workers do not respond more negatively to critical female supervisors. One possible mechanism to explain these differences is that Indian workers are less likely to hold gendered expectations and perceive the accuracy of criticism from women differently. Alternatively, stronger ideas of hierarchy may lead to more deference given to managers regardless of gender.

[20] Our results remain consistent when we include participants from Karnataka, Telangana, and Andhra Pradesh in the sample of southern workers.

These results present some hopeful evidence, at least for the narrow (but growing) part of the labor market we are studying. Future research should test whether results are similar in other parts of the labor market that are representative of other parts of the Indian economy, including more traditional features of worker-manager relationships such as more frequent and meaningful in-person and virtual interactions.

Appendix A

Tables

TABLE A1

FEEDBACK EFFECTS ON ATTITUDE BY WORKER GENDER (*N* = 600)

	Work Future		Job Satisfaction		Task Import		Attitude Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Female Workers							
Feedback	−.007	.014	−.069	−.143	−.013	−.075	−.035	−.041
	(.016)	(.025)	(.070)	(.107)	(.082)	(.120)	(.054)	(.083)
Above × feedback		−.042		.123		.110		−.002
		(.032)		(.138)		(.162)		(.107)
Above average		.030		.305***		.126		.212**
		(.024)		(.117)		(.136)		(.086)
Observations	600	600	600	600	600	600	600	600
	B. Male Workers							
Feedback	−.010	−.026*	−.074	−.116	−.082	−.172**	−.076**	−.145**
	(.009)	(.014)	(.048)	(.076)	(.054)	(.080)	(.036)	(.057)
Above × feedback		.030*		.034		.150		.105
		(.018)		(.093)		(.106)		(.071)
Above average		.004		.500***		.207**		.262***
		(.014)		(.073)		(.085)		(.055)
Observations	1,601	1,601	1,601	1,601	1,601	1,601	1,601	1,601
Sample mean	.96	.96	5.23	5.23	5.17	5.17	−.00	−.00
SD	.18	.18	.94	.94	1.05	1.05	.72	.72
p-value: feedback: M = F	.89		.95		.48		.52	
p-value: negative feedback: M = F		.16		.83		.50		.30
p-value: positive feedback: M = F		.15		.54		.66		.97

Note. Work future is a binary variable whether workers are interested to work for the firm in the future. Job satisfaction and task import measure how strongly workers agree that they were satisfied with the task and that the task was important, respectively (0 = strongly disagree, 6 = strongly agree). Attitude index is a standardized index of the other three outcome measures. *p*-Values indicate whether the effect of positive feedback is significantly different from zero. All estimations are by OLS. Robust standard errors are in parentheses. M = male; f = female.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

TABLE A2

FEEDBACK EFFECTS ON EFFORT BY WORKER GENDER

	Legible		Adding		Transcription Score		Voluntary Task	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Female Workers							
Feedback	.028	.047	.010	.023	.205	.173	.018	.049
	(.022)	(.039)	(.029)	(.043)	(.125)	(.180)	(.032)	(.047)
Above × feedback		−.035		−.029		.083		−.064
		(.045)		(.057)		(.249)		(.065)
Above average		−.017		.043		−.235		.075
		(.040)		(.053)		(.247)		(.060)
Observations	600	600	600	600	600	600	568	568
	B. Male Workers							
Feedback	.029**	.061***	.020	.035	.151**	.270***	.028	.042
	(.013)	(.022)	(.016)	(.023)	(.076)	(.101)	(.020)	(.029)
Above × feedback		−.068***		−.034		−.246		−.029
		(.026)		(.033)		(.150)		(.040)
Above average		.109***		.125***		.367**		.055
		(.027)		(.033)		(.148)		(.037)
Observations	1,601	1,601	1,601	1,601	1,601	1,601	1,496	1,496
Sample mean	.63	.63	.52	.52	12.83	12.83	.83	.83
SD	.34	.34	.40	.40	2.16	2.16	.37	.37
p-value: feedback: M = F	.97		.74		.71		.78	
p-value: negative feedback: M = F		.75		.80		.63		.91
p-value: positive feedback: M = F		.49		.87		.26		.58

Note. The dependent variable in cols. 1 and 2 is whether the worker says the receipt is legible. The dependent variable in cols. 3 and 4 measures whether the worker is willing to add up the amounts. The dependent variable in cols. 5 and 6 captures the accuracy of transcribing receipts. Columns 7 and 8 measure effects on the voluntary task. All estimations are by OLS. Robust standard errors are in parentheses. The *p*-values reported in the bottom three rows test whether responses to female and male managers are different by worker gender.

** $p < .05$.

*** $p < .01$.

TABLE A3

EFFECT OF FEEDBACK CONTENT ON ATTITUDES BY MANAGER GENDER (*N* = 2,201)

	Work Future		Job Satisfaction		Task Import		Attitude Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager female	−.007	−.024	−.021	−.103	.039	−.007	−.004	−.069
	(.011)	(.017)	(.063)	(.097)	(.071)	(.104)	(.046)	(.071)
Feedback	−.017	−.028*	−.091	−.150*	−.048	−.152	−.080*	−.148**
	(.011)	(.016)	(.056)	(.085)	(.066)	(.094)	(.043)	(.065)
Manager female × feedback	.016	.026	.036	.055	−.020	.031	.033	.070
	(.016)	(.024)	(.079)	(.124)	(.090)	(.134)	(.060)	(.094)
Above × feedback		.022		.074		.186		.112
		(.023)		(.107)		(.130)		(.084)
Female × feedback × above		−.021		−.033		−.104		−.074
		(.032)		(.154)		(.179)		(.118)

TABLE A3 (*Continued*)

	Work Future		Job Satisfaction		Task Import		Attitude Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample mean	.96	.96	5.24	5.24	5.18	5.18	.00	.00
SD	.19	.19	.91	.91	1.03	1.03	.70	.70

Note. Work future is a binary variable whether workers are interested to work for the firm in the future. Job satisfaction and task import measure how strongly workers agree that they were satisfied with the task and that the task was important, respectively (0 = strongly disagree, 6 = strongly agree). Attitude index is a standardized index of the three outcome measures. All estimations are by OLS. Robust standard errors are in parentheses.
* $p < .10$.
** $p < .05$.

TABLE A4

EFFECT OF FEEDBACK CONTENT ON EFFORT BY WORKER GENDER

	Legible		Adding		Transcription Score		Voluntary Task	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager female	.028	.021	.019	.006	.272	.117	−.012	−.014
	(.022)	(.032)	(.026)	(.033)	(.346)	(.455)	(.027)	(.039)
Feedback	.048**	.059**	.023	.010	.667**	.697	.008	.028
	(.021)	(.029)	(.024)	(.032)	(.325)	(.432)	(.025)	(.035)
Manager female × feedback	−.025	−.038	.006	.013	−.230	−.506	.036	.034
	(.028)	(.042)	(.034)	(.045)	(.444)	(.606)	(.034)	(.050)
Above × feedback		−.057		−.013		−.663		−.044
		(.036)		(.044)		(.560)		(.049)
Female × feedback × above		.042		.003		.770		.008
		(.051)		(.061)		(.769)		(.068)
Observations	2,201	2,201	2,201	2,201	2,201	2,201	2,064	2,064
Sample mean	.61	.61	.50	.50	8.99	8.99	.83	.83
SD	.34	.34	.39	.39	5.28	5.28	.37	.37

Note. The dependent variable in cols. 1 and 2 is whether the worker says the receipt is legible. The dependent variable in cols. 3 and 4 measure whether the worker is willing to add up the amounts. The dependent variable in cols. 5 and 6 captures the accuracy of transcribing receipts. Columns 7 and 8 use the score as the dependent variable and divide the sample by worker gender. All estimations are by OLS. Robust standard errors are in parentheses. The *p*-value reported in the last row tests whether the sum of the two feedback coefficients is significantly different from zero.
** $p < .05$.

TABLE A5

EFFECT OF FEEDBACK PERCEPTION ACCURACY AND APPROPRIATENESS OF FEEDBACK

	Feedback Accurate				Feedback Appropriate			
	Full (1)	Full (2)	Women (3)	Men (4)	Full (5)	Full (6)	Women (7)	Men (8)
Female manager	.000	.004	.213	−.094	.038	.023	.056	−.013
	(.072)	(.117)	(.207)	(.141)	(.065)	(.111)	(.210)	(.130)
Above × female manager		−.043	−.478*	.140		−.009	−.195	.078
		(.145)	(.272)	(.169)		(.131)	(.257)	(.152)
Above average		.606***	.799***	.527***		.625***	.637***	.612***
		(.105)	(.206)	(.120)		(.093)	(.180)	(.107)

TABLE A5 (*Continued*)

	Feedback Accurate				Feedback Appropriate			
	Full (1)	Full (2)	Women (3)	Men (4)	Full (5)	Full (6)	Women (7)	Men (8)
Observations	1,302	1,302	363	939	1,302	1,302	363	939
Sample mean	4.91	4.91	4.88	4.92	5.02	5.02	4.98	5.05
SD	1.303	1.303	1.219	1.314	1.181	1.181	1.247	1.168

Note. Dependent variables are how much workers agree that the feedback was accurate (cols. 1–4) and appropriate (cols. 5–8) on a scale from 0 (strongly disagree) to 6 (strongly agree). All estimations are by OLS. Robust standard errors are in parentheses. The table is estimated for participants who received feedback.

* $p < .10$.

*** $p < .01$.

References

Abbas, S. 2015. "Indian Organizations and Attitude Towards Women Managers: Gender Differential, Sex Bias, Reaction and Perceptive Reality." *Yobe Journal of Economics* 2:117–28.

Abel, M. 2024. "Do Workers Discriminate against Female Bosses?" *Journal of Human Resources* 59, no. 2:470–501.

Amba-Rao, S. C., J. A. Petrick, J. N. Gupta, and T. J. V. D. Embse. 2000. "Comparative Performance Appraisal Practices and Management Values among Foreign and Domestic Firms in India." *International Journal of Human Resource Management* 11, no. 1:60–89.

Asad, S. A., R. Banerjee, and J. Bhattacharya. 2020. "Do Workers Discriminate against Their Out-Group Employers? Evidence from the Gig Economy." IZA Discussion Paper no. 13012, Institute of Labor Economics, Bonn

Ayalew, S., S. Manian, and K. Sheth. 2021. "Discrimination from Below: Experimental Evidence from Ethiopia." *Journal of Development Economics* 151:102653.

Azmat, G., and N. Iriberri. 2010. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students." *Journal of Public Economics* 94, nos. 7–8:435–52.

Bandiera, O., V. Larcinese, and I. Rasul. 2015. "Blissful Ignorance? A Natural Experiment on the Effect of Feedback on Students' Performance." *Labour Economics* 34:13–25.

Barcellos, S. H., L. S. Carvalho, and A. Lleras-Muney. 2014. "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?" *American Economic Journal: Applied Economics* 6, no. 1:157–89.

Bartos, V., M. Bauer, J. Chytilová, and F. Matějka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review* 106, no. 6:1437–75.

Basu, S. 2016. "One in Two Indian Employees Prefer Telecommuting, Reveals Randstad Survey." *Economic Times*, June 9.

Batra, R., and T. G. Reio Jr. 2016. "Gender Inequality Issues in India." *Advances in Developing Human Resources* 18, no. 1:88–101.

BCG (Boston Consulting Group). 2021. "Unlocking the Potential of the Gig Economy in India." Boston, MA: Boston Consulting Group.

Beaman, L., E. Duflo, R. Pande, and P. Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science* 335, no. 6068:582–86.

Bertrand, M. 2011. "New Perspectives on Gender." In *Handbook of Labor Economics*, vol. 4B, edited by O. Ashenfelter and D. Card, 1543–90. Amsterdam: North-Holland.

Blanes i Vidal, J., and M. Nossol. 2011. "Tournaments without Prizes: Evidence from Personnel Records." *Management Science* 57, no. 10:1721–36.

Blau, F. D., and L. M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55, no. 3:789–865.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics* 128, no. 1:1–51.

Bloom, N., R. Sadun, and J. Van Reenen. 2012. "The Organization of Firms across Countries." *Quarterly Journal of Economics* 127, no. 4:1663–705.

Boring, A. 2016. "Gender Biases in Student Evaluations of Teaching." *Journal of Public Economics* 145:27–41.

Bracken, C. C., L. W. Jeffres, and K. A. Neuendorf. 2004. "Criticism or Praise? The Impact of Verbal versus Text-Only Computer Feedback on Social Presence, Intrinsic Motivation, and Recall." *Cyberpsychology and Behavior* 7, no. 3:349–57.

Budhwar, P. S., D. S. Saini, and J. Bhatnagar. 2005. "Women in Management in the New Economic Environment: The Case of India." *Asia Pacific Business Review* 11, no. 2:179–93.

Burgoon, J. K. 1993. "Interpersonal Expectations, Expectancy Violations, and Emotional Communication." *Journal of Language and Social Psychology* 12, nos. 1–2:30–48.

Buser, W., J. Hayter, and E. C. Marshall. 2019. "Gender Bias and Temporal Effects in Standard Evaluations of Teaching." *AEA Papers and Proceedings* 109:261–65.

Carli, L. L. 2001. "Gender and Social Influence." *Journal of Social Issues* 57, no. 4:725–41.

Catalyst. 2019. "Quick Take: Women in the Workforce—India." New York: Catalyst.

Cecchi-Dimeglio, P. 2017. "How Gender Bias Corrupts Performance Reviews, and What to Do about It." *Harvard Business Review*, April 12.

Chakraborty, P., and D. Serra. 2024. "Gender and Leadership in Organizations: The Threat of Backlash." *Economic Journal* 134, no. 660:1401–30.

Charness, G., D. Masclet, and M. C. Villeval. 2014. "The Dark Side of Competition for Status." *Management Science* 60:38–55.

Coffman, K. B., C. B. Flikkema, and O. Shurchkov. 2021. "Gender Stereotypes in Deliberation and Team Decisions." *Games and Economic Behavior* 129:329–49.

Correll, S., and C. Simard. 2016. "Vague Feedback Is Holding Women Back." *Harvard Business Review*, April 29.

Das, M. 2000. "Women Entrepreneurs from India: Problems, Motivations and Success Factors." *Journal of Small Business and Entrepreneurship* 15, no. 4:67–81.

Dhar, D., T. Jain, and S. Jayachandran. 2018. "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." NBER Working Paper no. 25331 (December), National Bureau of Economic Research, Cambridge, MA.

Eagly, A. H., S. J. Karau, and M. G. Makhijani. 1995. "Gender and the Effectiveness of Leaders: A Meta-analysis." *Psychological Bulletin* 117, no. 1:125.

Eriksson, T., A. Poulsen, and M. C. Villeval. 2009. "Feedback and Incentives: Experimental Evidence." *Labour Economics* 16, no. 6:679–88.

Fryer, R. G., Jr., and S. D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119, no. 3:767–805.

Gjedrem, W. G. 2018. "Relative Performance Feedback: Effective or Dismaying?" *Journal of Behavioral and Experimental Economics* 74:1–16.

Glover, D., A. Pallais, and W. Pariente. 2019. "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores." *Quarterly Journal of Economics* 132, 3:1219–60.

Goldin, C. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104, no. 4:1091–119.

Greenwald, A. G., M. R. Banaji, and B. A. Nosek. 2015. "Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects." *Journal of Personality and Social Psychology*108, no. 4:553–61.

Greenwald, A. G., B. A. Nosek, and M. R. Banaji. 2003. "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm." *Journal of Personality and Social Psychology* 85, no. 2:197.

Grossman, P. J., C. Eckel, M. Komai, and W. Zhan. 2016. "It Pays to Be a Man: Rewards for Leaders in a Coordination Game." Monash Economics Working Paper no. 38-16, Department of Economics, Monash University.

Gupta, A., M. Koshal, and R. Koshal. 1998. "Women Managers in India: Challenges and Opportunities." *Equal Opportunities International* 17, no. 8:14–26.

Gupta, N., L. Rigotti, and A. Wilson. 2021. "The Experimenters' Dilemma: Inferential Preferences over Populations." https://doi.org/10.48550/arXiv.2107.05064.

Hoffman, M., L. B. Kahn, and D. Li. 2017. "Discretion in Hiring." *Quarterly Journal of Economics* 133, no. 2:765–800.

Holmstrom, B., and P. Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7:24–52.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow. 2019. "The Allocation of Talent and US Economic Growth." *Econometrica* 87, no. 5:1439–74.

Ipsos. 2011. "Telecommuting: Citizens in 24 Countries Assess Working Remotely for a Total Global Perspective." Ipsos Global Advisor, November.

Jayachandran, S., and R. Pande. 2017. "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." *American Economic Review* 107, no. 9:2600–29.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75, no. 1:83–119.

Kuhnen, C. M., and A. Tymula. 2012. "Feedback, Self-Esteem, and Performance in Organizations." *Management Science* 58, no. 1:94–113.

Kwantes, C. T. 2009. "Culture, Job Satisfaction and Organizational Commitment in India and the United States." *Journal of Indian Business Research* 1, no. 4:196–212.

List, J. A., and F. Momeni. 2017. "When Corporate Social Responsibility Backfires: Theory and Evidence from a Natural Field Experiment." NBER Working Paper no. 24169 (December), National Bureau of Economic Research, Cambridge, MA.

Macchiavello, R., A. Menzel, A. Rabbani, and C. Woodruff. 2020. "Challenges of Change: An Experiment Promoting Women to Managerial Roles in the Bangladeshi Garment Sector." NBER Working Paper no. 27606 (July), National Bureau of Economic Research, Cambridge, MA.

Manne, K. 2017. *Down Girl: The Logic of Misogyny*. Oxford: Oxford University Press.

McKenzie, D., and C. Woodruff. 2015. "Business Practices in Small Firms in Developing Countries." *Management Science* 63, no. 9:2773–3145.

Mengel, F., J. Sauermann, and U. Zölitz. 2019. "Gender Bias in Teaching Evaluations." *Journal of the European Economic Association* 17, no. 2:535–66.

Niederle, M., and L. Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122, no. 3:1067–101.

Noble House. 2019. "Gig Workforce: The Future of Work Is Anywhere." Haryana, India: Noble House Consulting.

Oswald, F. L., G. Mitchell, H. Blanton, J. Jaccard, and P. E. Tetlock. 2013. "Predicting Ethnic and Racial Discrimination: A Meta-analysis of IAT Criterion Studies." *Journal of Personality and Social Psychology* 105, no. 2:171–92.

Podsakoff, P. M., S. B. MacKenzie, R. H. Moorman, and R. Fetter. 1990. "Transformational Leader Behaviors and Their Effects on Followers' Trust in Leader, Satisfaction, and Organizational Citizenship Behaviors." *Leadership Quarterly* 1, no. 2:107–42.

Rammohan, A., and P. Vu. 2018. "Gender Inequality in Education and Kinship Norms in India." *Feminist Economics* 24, no. 1:142–67.

Ranganathan, A., and R. Shivaram. 2021. "Getting Their Hands Dirty: How Female Managers Motivate Female Worker Productivity through Subordinate Scut Work." *Management Science* 67, no. 5:2657–3320.

Reuben, E., P. Sapienza, and L. Zingales. 2014. "How Stereotypes Impair Women's Careers in Science." *Proceedings of the National Academy of Sciences USA* 111, no. 12:4403–8.

Sully De Luque, M. F., and S. M. Sommer. 2000. "The Impact of Culture on Feedback-Seeking Behavior: An Integrated Model and Propositions." *Academy of Management Review* 25, no. 4:829–49.

Varma, A., S. Pichler, and E. S. Srinivas. 2005. "The Role of Interpersonal Affect in Performance Appraisal: Evidence from Two Samples—the US and India." *International Journal of Human Resource Management* 16, no. 11:2029–44.

Woetzel, L., A. Madgavkar, K. Sneader, O. Tonby, D.-Y. Lin, J. Lydon, S. Sha, M. Krishnan, K. Ellingrud, and M. Gubieski. 2018. "The Power of Parity: Advancing Women's Equality in Asia Pacific." Washington, DC: McKinsey Global Institute.

Zimmermann, L. 2012. "Reconsidering Gender Bias in Intrahousehold Allocation in India." *Journal of Development Studies* 48, no. 1:151–63.