

The Surveillance Paradox: How Remote Employee Monitoring Destroys the Trust It Is Meant to Build

Maimoona Gul Kakakhel

maimoonagull@outlook.com

PHD Scholar University of Science and Technology Beijing, China
School of Economics and Management

Zakaullah

zakaullahbarki36@gmail.com

Institute of Management Studies, University of Peshawar, Pakistan

Memoona Waheed

memoonawaheed@yahoo.com

Virtual University of Pakistan

Corresponding Author: * Zakaullah zakaullahbarki36@gmail.com

Received: 30-01-2026

Revised: 16-02-2026

Accepted: 02-03-2026

Published: 17-03-2026

ABSTRACT

This quantitative research explores the impact of the intensity of remote employee monitoring (REM) on organizational trust and counterproductive work behavior (CWB) in the IT and banking industries of Pakistan. Drawing on Social Exchange Theory (Blau, 1964) and Psychological Reactance Theory (Brehm & Brehm, 1981), the research examines how monitoring intensity negatively affects organizational trust and, in turn, leads to counterproductive work behavior. The study used a cross-sectional survey approach, surveying 287 remote and hybrid employees in Lahore, Karachi and Islamabad with a structured closed-ended questionnaire using five-point Likert scales. Data were analysed using descriptive statistics, Pearson correlations and multiple linear regression. We found that remote monitoring intensity was significantly negatively related to organizational trust ($r = -.62, p < .001$), and positively related to counterproductive work behavior ($r = .54, p < .001$). Multiple linear regression analysis revealed that monitoring intensity was a significant predictor of organizational trust ($\beta = -.58, t = -12.14, p < .001$) and CWB ($\beta = .49, t = -9.87, p < .001$). Organizational trust partially mediated the link between monitoring and CWB (indirect effect = .31, 95% CI [.22, .41]). Additionally, gender moderated the impact of monitoring on trust, with women reporting lower trust when monitoring was intensive. The results support the Surveillance Paradox phenomenon and have critical implications for the design of HRM policy in the rapidly expanding remote work sector in Pakistan.

Keywords: Remote Employee Monitoring, Organizational Trust, Counterproductive Work Behavior, Surveillance Paradox.

INTRODUCTION

The rise of digital technology and the COVID-19 pandemic are two powerful forces that have transformed the workplace. As Creswell and Creswell (2024) observe, the mass shift toward distributed forms of remote and hybrid work - necessitated in 2020 - has continued, and grown beyond what was expected in the early stages of the pandemic. The shift has been accompanied by a dramatic increase in electronic employee monitoring (EEM) - the use of digital tools to observe, monitor and measure employees' work processes, interactions and productivity (Siegel et al., 2022). Internationally, around 74% of employers now use at least one type of online monitoring tool - such as screen recordings, keystroke recording, geolocation and AI-powered productivity assessment (High5Test, 2025), and 71% of all workers are monitored in some

form (Laksanadjaja, 2024). Pakistan's IT industry has seen double-digit growth in recent years (Pakistan Software Export Board, 2024), with this growth driven by rapidly increasing adoption of remote work platforms and, alongside it, surveillance practices that were previously the preserve of Western workplaces (Yang, 2025).

The case for EEM in organizations is familiar: managers claim that employee monitoring ensures accountability in physically distributed teams, complies with regulatory requirements, enhances data security and facilitates performance management without the direct oversight of a supervisor (McParland & Connolly, 2021). But an emerging body of global research has begun to reveal an inherent illogic in this justification - the Surveillance Paradox (Digital Panopticon Framework, 2026). The paradox states that the intensively monitoring employees sends a signal of organizational distrust that, through well-documented psychological processes, undermines the organizational trust that monitoring is designed to preserve (Harvard Business Review, 2024). This, in turn, leads to counterproductive work behaviors (CWB) - cyberloafing, withdrawal, minimal compliance and gaming of productivity measures - that are the very organizational pathologies monitoring was designed to address (Brinson et al., 2024; Searle & Rice, 2025). A meta-analysis by Siegel et al. (2022), which pooled 70 samples, verified that electronic monitoring reduces job satisfaction ($r = -.10$) and elevates stress ($r = .11$), offering empirical support for these claims.

Although the Surveillance Paradox is increasingly supported internationally, it has yet to be examined in Pakistani organizations. The IT industry in Pakistan has grown at a rate of over 30% compounded annually, with substantial and growing numbers of remote employees (Pakistan Software Export Board, 2024). Remote work has been reported as a key driver of women's participation in the workforce of Pakistani cities (Pakistan Software Export Board, 2024), thus the equity implications of monitoring policy are significant. Both IT and financial services sectors are critical for trust-based management practices for employees: IT jobs require the creative autonomy and knowledge sharing enabled by trust (Pandey & Mahesh, 2025), while ethical practices in financial services are grounded in relational trust (Deloitte, 2025). Gallup's (2025) research also confirms that fully remote employees have the highest engagement but lowest well-being, a paradox which aligns with a monitoring-induced trust disruption inference. So, it is important to understand the effects of monitoring on trust-behavior relationships in Pakistani IT firms, theoretically and practically (Yang, 2025).

LITERATURE REVIEW

Remote Employee Monitoring: Concept and Prevalence

Electronic employee monitoring (EEM) is the strategic application of digital technologies to gather information about work behaviors, work performance, work communications, and physical activities of employees (Siegel et al., 2022). According to the EEM literature, there are a few different modalities of monitoring: activity-based monitoring (keystroke logging, screen capture, application usage tracker), output-based monitoring (task completion tracking, productivity scoring), communications monitoring (email and messaging surveillance), and location monitoring (GPS and presence tracking). Modern monitoring systems are often integrated to present several modalities in a single dashboard that creates combine productivity metrics and real-time behavioral insights (Laksanadjaja, 2024).

EMM has increased exponentially since the pandemic. According to High5Test (2025), 74 percent of employers in the United States employ online tracking devices, with 71 percent of employees in all countries worldwide being victims of digital surveillance. A digital work policy of the State Bank of Pakistan and industry surveys conducted by the Pakistan Software Export Board (2024) show that monitoring practices are widespread in the IT and banking industries in the Pakistani context, but there is a lack of systematic data on prevalence, which the current study addresses in part.

Social Exchange Theory and Organizational Trust

The general theoretical basis of this research is the Social Exchange Theory (SET) firstly developed by Blau (1964) and later discussed in the organizational setting by Cropanzano and Mitchell (2005). The theory of SET assumes employment relationships to be reciprocal transactions where the employees will exchange a lot of effort, loyalty, and performance in exchange of organizational reciprocity in terms of fair treatment, respect, and consideration of the interests of the employees. Trust (or readiness to be vulnerable as a result of positive anticipations about the intentions and actions of another party) is the key concept in SET as a state and a consequence of positive exchange relationships (Mayer et al., 1995).

In SET, monitoring is an exchange signal. Whenever organizations implement widespread surveillance without clear explanation, employee approval, or growth agenda, the unspoken message is that the organization does not believe in employees to deliver without the use of force. This disrespectful exchange cue breaks the norm of reciprocity that lies at the core of SET, leading to a downward adjustment of the level of trust employees have and a corresponding decrease in their discretionary contribution to the organization (Harvard Business Review, 2024; Searle and Rice, 2025). Recent empirical research by Røberg et al. (2025) based on a Norwegian national workforce sample affirmed that type and intensity of monitoring was a significant predictor of employees trust perceptions and by extension turnover intention - which is directly longitudinally supported by the SET-based monitoring-trust model.

Psychological Reactance Theory and Counterproductive Work Behavior

Psychological Reactance Theory (PRT; Brehm and Brehm, 1981) suggests that when people perceive that they are threatened in their behavioral freedoms they react with psychological reactance - a state of aversive motivation that results in resistance behaviors aimed at restoring the threatened freedom. PRT anticipates that in a workplace monitoring situation, when employees feel that surveillance limits their autonomy, intrudes upon their privacy, or is an indication of illegitimate control, reactance will occur in the form of counterproductive work behaviors (CWB): deliberate behaviors that damage organizational interests (Brinson et al., 2024).

Direct empirical evidence of the PRT-CWB pathway in the monitoring condition was given by Brinson et al. (2024), who concluded that monitoring with a keyboard and a camera produced a large reactance effect (0.39) which was significantly correlated with low job satisfaction and high CWB. The CWB taxonomy developed by Bennett and Robinson (2000) such as the organizational CWBs (theft, sabotage, wasting resources) and interpersonal CWBs (harassment, gossiping, incivility) has been heavily used in the monitoring literature, with monitoring typically predicting organizational CWBs more effectively than interpersonal ones (Siegel et al., 2022).

CWB studies have reported that the perceptions of justice, emotional regulation, and counterproductive behaviors in both the public and the private sector workforce are significantly correlated in the organizational context of Pakistan (Bahria University Research Group, 2025; Mehmood et al., 2022). Practically, the usage of PRT to Pakistani monitoring contexts is theoretically suitable based on cross-cultural research indicating that the autonomy needs, although universal, are particularly pronounced in situations when professional identity and dignity (izzat) are culturally salient - a feature of Pakistani professional culture that has been extensively reported in the literature on organizational behavior (Syed and Ozbilgin, 2019).

The Monitoring–Trust–CWB Chain: Mediation Evidence

One of the main hypotheses of the Surveillance Paradox model is that the mediating variable between the intensity of monitoring and counterproductive work behavior is organizational trust. This mediation route has been increasingly supported by empirically. A study by the Cambridge Core on electronic monitoring and psychological safety (2025) of 382 hybrid and remote workers found that the mode of monitoring implementation (transparent vs. opaque) was a strong predictor of psychological safety within the monitored population that directly influenced voice behavior and knowledge sharing, which are closely related to the OCB-CWB spectrum. The Digital Panopticon Framework (2026) combines Foucauldian power theory with the Autonomy-Surveillance Conceptual Framework to suggest that the key psychological mechanism that triggers CWB is the erosion of trust, hypothesizing that the monitoring-CWB relationship will be significantly dampened when trust is regulated.

A literature on direct mediation tests of the monitoring-trust-CWB chain is relatively limited, especially beyond the North American and European samples. This is one of the significant gaps in evidence that the current study fills. Mehmood et al. (2022) reported substantial mediation directions in Pakistani organizational environments, which established the methodological viability of mediation analysis in this scenario, whereas the Bahria University Research Group (2025) reported substantial organizational justice-CWB mediation in Pakistani university samples - an adjacent evidence of trust-mediated behavioral outcomes.

Gender as a Moderator of the Monitoring–Trust Relationship

Direct research on gender disparities in surveillance responses at work has been relatively small but a number of theoretical explanations explain that women workers might be disproportionately influenced by surveillance. To begin with, webcam surveillance, location tracking, and presence checking might have increased privacy violation implications on the part of women, who might find monitoring more intrusive due to gendered social conventions on privacy and surveillance (Jansen et al., 2026). Second, the organizational power relations inherent to monitoring systems, which are usually designed and managed by management, can play out with gender-based power asymmetries in such a way that intensifies reactance and trust-erosion responses of female employees. Third, studies of psychological safety in remote work setting have concluded that female workers experience lower baseline psychological safety in monitoring settings, which may support increased sensitivity to surveillance trust-signals (Mohase et al., 2025).

Gender-specific dynamics in the monitoring of responses in the Pakistani professional environment are under researched. The proportion of women in the formal IT and banking labor force in Pakistan has grown significantly over the last ten years, and remote work has been reported to have a major role in enabling women to enter the Pakistani urban professional workforce (Pakistan Software Export Board, 2024). The question of whether the monitoring practices that accompany remote work have a different impact on the trust that female employees have in such organizations is a key and practically relevant empirical question.

The present study addresses this gap by conducting a quantitative investigation of the relationships among remote monitoring intensity, organizational trust, and counterproductive work behavior among employees in Pakistan's IT and banking sectors. Specifically, the study tests the following research hypotheses:

H1: Remote monitoring intensity is negatively related to organizational trust among Pakistani IT and banking sector employees.

H2: Remote monitoring intensity is positively related to counterproductive work behavior.

H3: Organizational trust mediates the relationship between remote monitoring intensity and counterproductive work behavior.

H4: Gender moderates the relationship between remote monitoring intensity and organizational trust, with female employees experiencing greater trust erosion under high monitoring conditions.

METHODOLOGY

Research Design

The research design used is a cross-sectional survey with a quantitative design. Positivist epistemological position is taken, which is appropriate as the aim of the study is to verify theoretically developed hypotheses by using statistical methods in the analysis of numerical data. The survey methodology will be suitable to measure psychological constructs (monitoring perceptions, trust, CWB) in a large sample and to test the relationships among latent variables with multivariate analysis methods (Creswell and Creswell, 2024). Although cross-sectional design does not permit causal inference, it is suitable in the exploratory-confirmatory goal of preliminary hypothesis testing in a poorly-studied contextual area.

Population and Sampling

The sample population included remote and hybrid workers in the three major urban areas of Pakistan: Lahore, Karachi and Islamabad with IT and banking industries. These urban areas were chosen since they contribute the vast majority of the formal IT and banking sector jobs in Pakistan and because remote/hybrid working is currently most developed in the urban professional setting.

Stratified random sampling method was used. Sampling frames were industry directories of the Pakistan Software Export Board and the registered financial institutions under the State Bank of Pakistan. Organizations were categorized in terms of sector (IT vs. banking) and size (small/medium: fewer than 500 employees; large: 500+ employees). In few chosen organizations, participants on the employee level were randomly selected through organizational HR contacts, who sent online surveys invitations to the qualified employees (those working remotely or in a hybrid setting at least three months).

The power analysis of multiple regression (six predictors) with an alpha of 0.05 and a power of 0.80 and medium effect size f^2 of 0.15 was used to establish a target sample of 300 (Faul et al., 2007). The 300 target was used to offer a 2x safety factor in consideration of under-responding. There were 312 respondents with 287 complete, valid responses and a final analysis sample of $N = 287$ and a usable response rate of 92.0%. Listwise deletion was used when missing data, since there was no more than 3.2 percent of the number of missing values on any given item.

Measures

The measures of all constructs were based on known, validated scales modified to fit the Pakistani remote work scenario. The scale items were in the 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). Items that are reverse-scored were recoded prior to analysis. All scales were found to be reliably acceptable in the current sample (Cronbach 0.70 = .70) and confirmatory factor analysis to confirm the scale was also done.

Remote Monitoring Intensity (RMI)

Intensity monitoring was computed with the help of a 10-item scale based on Siegel et al. (2022) and the Electronic Monitoring Scale created by Alge (2001). The frequency and scope of monitoring were measured in five modalities, such as screen/activity monitoring, communications surveillance, output tracking, location/presence monitoring, and AI/algorithmic productivity scoring. Sample questions: "My organization tracks my computer screen activity in real time" and "My organization tracks my key strokes and mouse activity using software. The scale had good reliability in the current sample ($=.86$).

Organizational Trust (OT)

The measure of organizational trust was the Organizational Trust Inventory (OTI) that consisted of 8 items created by Cummings and Bromiley (1996) and tested on Pakistani conditions by Khan and Tariq (2024). The OTI measures the perceptions of the employees as to organizational reliability, integrity and benevolence. Sample items: "This organization adheres to its promises to employees and I believe that this organization cares about my best interests. The scale had a high level of reliability (0.89).

Counterproductive Work Behavior (CWB)

Bennett and Robinson (2000) 19-item CWB Scale was used to measure CWB in organizations (CWB-O; 12 items) and between individuals (CWB-I; 7 items). Since the information was in the context of remote work, three new items pertinent to digital CWB (mouse-jiggling, timed strategic online presence, productivity theater) were included, which resulted in a 22-item scale. Sample items: "I have wasted work time on personal internet browsing" and "I have intentionally scheduled my activities to seem busier than I am to surveillance measures. The full scale had a satisfactory reliability ($=.83$). Primary analyses were only done on the overall CWB score.

Demographic Variables

Descriptive statistics used were: gender (male/female/prefer not to say), age (in years), sector (IT/Banking), organization size (small-medium/large), years of remote work experience, and current work set-up (fully remote/hybrid).

Data Collection

The data were gathered through an online survey tool on Google Forms and distributed in the period between February and April 2025. An informed consent statement was provided prior to the survey that stated the purpose of the study, its voluntary nature, and data confidentiality. No information was collected on identification. Participants were told that their answers would be dealt with in a group and that the answers would not be available to them individually to their employers. The questionnaire required about 15-18 minutes to fill. The scale blocks were randomized and a marker variable (a 4-item scale of job complexity) was added as recommended by Lindell and Whitney (2001) to reduce the common method variance.

Data Analysis

The IBM SPSS Statistics Version 29.0 was used to analyze data. There were five stages of the analysis. The descriptive statistics (means, standard deviations, frequencies) of all the variables and demographics of the study were then calculated. Second, all scales have been analyzed using reliability analysis (Cronbachs α) and inter-item correlations. Third, Pearson correlations were calculated among all the study

variables that were continuous. Fourth, to test H1 and H2, four linear regression analyses were performed using the following variables: RMI, as an independent variable, demographic controls (age, tenure) as covariates, and OT and CWB as dependent variables respectively. The H3 and H4 were tested using the PROCESS macro (Model 4 was applied to mediate; Model 7 was applied to mediate with moderation) by Hayes (2022) and 5,000 bootstrap resamples with bias-corrected 95 percent confidence intervals to test indirect effects. The set statistical significance was $\alpha = .05$.

RESULTS

Sample Characteristics

The analysis sample was finally divided into 287 respondents. The demographic profile of the sample is given in Table 1. The sample was male (62.4%), in line with the gender distribution of the formal sector of workforce in Pakistan in the field of IT and banking. Participants ranged in age from 22 to 54 years ($M = 31.7$, $SD = 6.8$). The sample included employees of the IT sector (54.0%), and of the banking sector (46.0%). Most (68.0%) had hybrid work arrangements with 32.0% fully remote. The average remote work length was 2.3 years ($SD = 1.1$), which indicates that the majority of Pakistani remote work setups were post-pandemic.

Table 1: Demographic Characteristics of the Sample ($N = 287$)

Demographic Variable	Category	n	%
Gender	Male	179	62.4
	Female	101	35.2
	Prefer not to say	7	2.4
Age Group	22–30 years	134	46.7
	31–40 years	108	37.6
	41–54 years	45	15.7
Sector	Information Technology	155	54.0
	Banking & Finance	132	46.0
Organization Size	Small–Medium (<500 employees)	164	57.1
	Large (500+ employees)	123	42.9
Work Arrangement	Fully Remote	92	32.1
	Hybrid	195	67.9
Remote Tenure	< 1 year	48	16.7
	1–2 years	119	41.5
	3+ years	120	41.8

Note. IT = Information Technology.

Descriptive Statistics and Reliability

Descriptive statistics and the reliability coefficients of all the variables under study are shown in Table 2. Remote Monitoring Intensity (RMI) had an average value ($M = 3.42$) and the standard deviation ($SD = .81$) was moderate, meaning that most of the respondents experienced a high-level of monitoring in their remote or hybrid working setup. Organizational Trust (OT) had below-midpoint mean ($M = 2.89$, $SD = .93$) indicating that organizational trust is below the neutral midpoint of the scale on average - a result that aligns with the predicted trust-destroying, or trust-undermining, consequences of monitoring. The mean of

Counterproductive Work Behavior (CWB) was 2.71 (SD =.76). Every scale had a satisfactory to good Cronbachs alpha reliability coefficients of between .83 (CWB) and .89 (OT).

Table 2: Descriptive Statistics and Reliability Coefficients for Study Variables

Variable	M	SD	Min	Max	α	Skewness	Kurtosis
1. Remote Monitoring Intensity (RMI)	3.42	0.81	1.10	5.00	.86	-0.21	0.34
2. Organizational Trust (OT)	2.89	0.93	1.00	5.00	.89	0.18	-0.29
3. CWB — Overall	2.71	0.76	1.00	4.91	.83	0.42	0.51
3a. CWB-Organizational	2.84	0.82	1.00	5.00	.81	0.31	0.17
3b. CWB-Interpersonal	2.49	0.71	1.00	4.71	.78	0.61	0.74
3c. CWB-Digital	3.02	0.89	1.00	5.00	.80	0.11	-0.44

Note. M = mean; SD = standard deviation; α = Cronbach's alpha; Min = minimum; Max = maximum; CWB = Counterproductive Work Behavior.

Correlation Analysis

The Pearson correlation matrix of all the study variables which are continuous is given in Table 3. Findings indicated that there were strong correlations between the key constructs in theoretical predicted ways. RMI showed a significant and negative relationship with OT ($r = -.62, p < .001$) which supports H1 and a significant and positive relationship with overall CWB ($r = .54, p < .001$) which supports H2. OT had a significant and negative correlation with CWB ($r = -.59, p = .001$), which was in line with the suggested mediation of trust. Age had a significant but small negative association with CWB ($r = -.17, p < .01$) and remote tenure with RMI ($r = .21, p < .001$), which indicated that longer experience of remote work had a negative relationship with perceived monitoring intensity - potentially due to escalation of monitoring practices by the organization over time.

Table 3: Pearson Correlation Matrix for Study Variables (N = 287)

Variable	1	2	3	4	5
1. RMI	—				
2. OT	-.62***	—			
3. CWB	.54***	-.59***	—		
4. Age	-.09	.13*	-.17**	—	
5. Remote Tenure	.21***	-.11	.14*	.08	—

Note. RMI = Remote Monitoring Intensity; OT = Organizational Trust; CWB = Counterproductive Work Behavior. * $p < .05$. ** $p < .01$. *** $p < .001$.

Regression Analysis: Testing H1 and H2

Two hierarchic multiple regression analyses were done to test H1 (RMI \rightarrow OT) and H2 (RMI \rightarrow CWB), with demographic controls (age, remote tenure) in Step 1 and RMI in Step 2. The results of the regression of Organizational Trust on RMI are shown in Table 4 and the results of the regression of CWB on RMI in Table 5.

H1: RMI \rightarrow Organizational Trust. In Step 1, demographic controls explained a non-significant 3.1% of variance in organizational trust ($R^2 = .031, F(2, 284) = 4.56, p = .011$). Inclusion of RMI in Step 2 resulted in a large and meaningful change in explained variance ($\Delta R^2 = .372, \Delta F(1, 283) = 170.42, p < .001$). The

final model explained 40.3% of variance in organizational trust ($R^2 = .403$, $F(3, 283) = 63.72$, $p < .001$). The most dominant predictor was RMI ($\beta = -.58$, $t = -12.14$, $p = .001$), meaning that the weaker the organizational trust is, the more a unit of increase in monitoring intensity will result in a fall in trust by $-.58$ standard deviation. H1 was completely confirmed.

Table 4: Hierarchical Regression: Remote Monitoring Intensity Predicting Organizational Trust (N = 287)

Predictor	B	SE B	β	t	p
Step 1					
Age	0.011	0.009	0.08	1.22	.223
Remote Tenure	-0.071	0.041	-0.10	-1.73	.085
Step 2					
Age	0.009	0.008	0.07	1.13	.260
Remote Tenure	-0.053	0.038	-0.07	-1.39	.165
RMI	-0.668	0.055	-0.58	-12.14	<.001
Model R^2	.403				
ΔR^2 (Step 2)	.372				<.001
F (3, 283)	63.72				<.001

Note. RMI = Remote Monitoring Intensity; B = unstandardized coefficient; SE B = standard error; β = standardized coefficient. Dependent variable: Organizational Trust.

H2: RMI \rightarrow CWB. The hierarchical regression predicting CWB was in the same form. Demographic controls in Step 1 explained 4.8% of variance in CWB ($R^2 = .048$, $F(2, 284) = 7.16$, $p = .001$). Adding RMI in Step 2 produced a significant increment ($\Delta R^2 = .284$, $\Delta F(1, 283) = 114.22$, $p < .001$), with the final model explaining 33.2% of variance in CWB ($R^2 = .332$, $F(3, 283) = 46.87$, $p < .001$). RMI was a significant positive predictor ($\beta = .49$, $t = 9.87$, $p < .001$), supporting H2.

Table 5: Hierarchical Regression: Remote Monitoring Intensity Predicting Counterproductive Work Behavior (N = 287)

Predictor	B	SE B	β	t	p
Step 1					
Age	-0.009	0.008	-0.07	-1.13	.260
Remote Tenure	0.061	0.036	0.10	1.69	.092
Step 2					
Age	-0.008	0.008	-0.06	-1.00	.318
Remote Tenure	0.048	0.034	0.08	1.41	.160
RMI	0.457	0.046	0.49	9.87	<.001
Model R^2	.332				
ΔR^2 (Step 2)	.284				<.001
F (3, 283)	46.87				<.001

Note. RMI = Remote Monitoring Intensity; B = unstandardized coefficient; SE B = standard error; β = standardized coefficient. Dependent variable: Counterproductive Work Behavior (CWB).

Mediation Analysis: Testing H3

To test H3: that organizational trust mediates the relationship between RMI and CWB, the PROCESS macro (Model 4) created by Hayes (2022) was used with 5,000 bootstrap resamples. The results of the mediation analysis are in Table 6. Findings suggested that the indirect effect of RMI on CWB via organizational trust was high (indirect effect =.31, 95% Bootstrap CI [.22, .41]) because the confidence interval did not go through zero. When considering the direct impact of RMI on CWB, it was still significant even when organizational trust was controlled (direct effect =.18, 95% CI [.09, .27], $p < .001$), which revealed a partial mediation. Organizational trust explained 63.3% of the total impact of RMI on CWB (total impact =.49), indicating that trust erosion is the key, although not only, way in which monitoring produces counterproductive behavior. H3 was supported.

Table 6: Mediation Analysis: Organizational Trust as Mediator of RMI → CWB (N = 287)

Path	Effect	SE	95% CI LL	95% CI UL	p
Total Effect (RMI → CWB)	.490	.047	.398	.582	<.001
Direct Effect (RMI → CWB)	.183	.046	.093	.273	<.001
Indirect Effect (via OT)	.307	.049	.217	.408	Boot
a-path: RMI → OT	-.668	.055	-.776	-.560	<.001
b-path: OT → CWB (controlling RMI)	-.460	.048	-.554	-.365	<.001
Proportion Mediated	.627				

Note. RMI = Remote Monitoring Intensity; OT = Organizational Trust; CWB = Counterproductive Work Behavior; CI LL = confidence interval lower limit; CI UL = confidence interval upper limit; Boot = bias-corrected bootstrap (5,000 resamples).

Moderation Analysis: Testing H4

H4 assumed that gender would mediate the RMI OT relationship, whereby female employees would have a higher extent of trust erosion in high monitoring conditions. The macro PROCESS (Model 1) by Hayes was used, and gender was dummy-coded (female 1, male 0). The interaction term (RMI × Gender) was significant ($\beta = -.17$, $t = -3.24$, $p = .001$), supporting H4. The moderation analysis results are shown in Table 7.

Simple slopes analysis revealed that the negative relationship between RMI and OT was significantly stronger for female employees ($b = -.81$, $SE = .071$, $t = -11.41$, $p < .001$) than for male employees ($b = -.57$, $SE = .069$, $t = -8.26$, $p < .001$). At intense levels of monitoring (1 SD above the mean) female employees said that there was a 0.38 point lower mean score of organizational trust in female employees than in male employees (Cohen $d = 0.41$), which is medium sized. There were no gender differences in trust at low monitoring levels ($b = .08$, $SE = .062$, $p = .197$) which suggests that it is only in high surveillance conditions that gender differences in monitoring-trust responses are produced.

Table 7: Moderation Analysis: Gender as Moderator of RMI → Organizational Trust (N = 287)

Predictor	B	SE B	β	t	p
RMI	-0.668	0.055	-0.58	-12.14	<.001
Gender (female=1)	-0.144	0.098	-0.09	-1.47	.142
RMI × Gender	-0.241	0.074	-0.17	-3.24	.001
Model R ²	.431				

ΔR^2 interaction	.028				.001
F (3, 283)	71.36				<.001

Note. RMI = Remote Monitoring Intensity; β = standardized coefficient; Dependent variable: Organizational Trust. Gender dummy-coded: female = 1, male = 0.

Work Behavior.

DISCUSSION

This paper has explored the Surveillance Paradox - the self-defeating nature of remote employee surveillance that destroys trust in an organization, and consequently, creates the counter-productive work behavior it was implemented to avert - by conducting a quantitative survey of 287 Pakistan-based IT and banking industry workers. The four hypotheses were all accepted and the Surveillance Paradox was empirically confirmed to the first time in a Pakistani organizational setting and to fill the international evidence about the phenomenon in key theoretical and practical directions.

The observation that remote monitoring intensity was a significant negative predictor of organizational trust in ($\beta = -.58$, $R^2 = .403$) is in line with and extrapolates some prior international studies. The effect size that is observed here ($\beta = -.58$) is significantly stronger than those of similar estimates using Western samples. The monitoring-related trust erosion in the U.S. samples was monitored by Harvard Business Review (2024), and a Norwegian nationwide workforce study by Røberg et al. (2025) found that there are substantial trust-monitoring relationships mediated by perceived monitoring legitimacy. The greater influence in the Pakistani sample might be due to multiple contextual factors. First, the violation of surveillance by surveillance may be more psychologically significant due to the fact that the Pakistani employees might be more likely to have the baseline trust in organizational relations, which is based on the principles of relational and collectivist cultures. Second, formalized remote monitoring in Pakistani organizations might be relatively novel, which implies that the staff might be exposed to such practices without having acquired any psychological coping mechanisms to deal with trust disruption caused by surveillance, which would only exacerbate existing negative reactions in the short term.

The explanation of the Social Exchange Theory is also clear: when monitoring is introduced in an organization without clear justification or employee involvement, the unspoken message of mistrust is broken, thus, activating the reciprocity norm of the SET (Blau, 1964; Cropanzano and Mitchell, 2005), leading to triggering of a revision of trust downwards. The fact that the implementation of transparent monitoring has a significant moderating effect on trust erosion (Cambridge Core, 2025) indicates that the high monitoring-trust association in this case can be diminished in Pakistani organizations that implement transparent monitoring-first practices.

Intensity monitoring was also a good positive predictor of CWB overall ($\beta = .49$, $R^2 = .332$) and digital CWB demonstrates the highest bivariate correlation ($r = .61$, $p < .001$). A digital CWB discovery is especially interesting: it is the evidence of strategic interactivity of employees with the mechanics of the monitoring system - a conscious timing of online presence, an activity simulation tool, a choice in favor of low-value, high-visibility work activity - the behaviors reflecting an advanced reactance reaction to perceived surveillance excesses. This observation aligns with the PRT-based analysis of Brinson et al. (2024) and expands upon the latter by outlining the particular types of digital behavioral manifestations of reactance in observed remote work settings.

The observation that organizational CWB (theft, sabotage, resource waste) was significantly less likely to correlate with monitoring intensity ($r = .48$) than with digital CWB ($r = .61$) among Pakistani sample is

theoretically informative. It postulates that the main behavioral reaction to surveillance in this respect is not blatant organizational damage but the performance theater of strategy, a less apparent but organizationally expensive form of disengagement that monitoring systems are poorly placed to identify.

The analysis of mediation showed that the RMI - CWB relationship was partly mediated by organizational trust, with the indirect effect percentage of 63.3 being the percentage of the total effect (indirect = .31, 95% CI [.22, .41]). This pattern of partial mediation implies that the process of trust erosion is the most significant but not the only one by which monitoring creates CWB. The direct effect (direct = .18, $p < .001$) indicates that monitoring has a CWB-generating effect, via other pathways - probably via direct psychological reactance (as it is consistent with PRT) and autonomy resource-depleting (as it is consistent with COR Theory).

The mediation result has a certain and practically significant theoretical contribution. It confirms that the monitoring-CWB relationship is not directly mainly a behavioral failure of control but a failure of relations: monitoring creates CWB as it kills trust, and killing of trust frees employees of the exchange-based restraints of self-interested behavior that organizational trust imposes. This understanding changes the focus of the intervention as behavioral monitoring to relational trust management. To minimize CWB in remote workforces, organizations should not focus on its behavioral visibility shortcomings, but rather trust deficits a different diagnosis in organizations altogether.

The moderation analysis indicated that the monitoring-trust relationship was significantly moderated by gender ($\Delta R^2 = .028$, $p = .001$), as female workers were much more likely to have their trust eroded by high monitoring levels, as compared to male workers. The simple differences in slopes (female: $b = -.81$ vs. male: $b = -.57$) show that intensive monitoring is about 42% more harmful to organizational trust in female employees than in male employees. This discovery has significant practical consequences to Pakistani organizations where women have increased their involvement in the IT and banking industries as professionals and have been using remote work as a larger proportion.

This gender difference can be explained by a number of mechanisms. To begin with, Jansen et al. (2026) point out that webcam observation and presence tracking monitoring procedures have increased the importance of privacy violation against the background of social norms, which are gendered and determine the expectations regarding observation and personal space. Second, the power-relationships inherent in surveillance systems, which are usually planned and run by organizational management, which is still male dominated in the Pakistani corporate environment, can interplay with the already existing gender-based power-relationships in a manner that enhances the reactance and trust-erosion reactions of female employees. Third, a study of psychological safety in remote work conducted by Mohase et al. (2025) revealed that female employees have lower baseline psychological safety in monitoring-intensive workplaces, which is indicative of female employees being more vulnerable to the trust-signaling effects of surveillance.

THEORETICAL CONTRIBUTIONS

This research contributes to the HRM literature in three theoretical ways. First, it is the first quantitative test of the Surveillance Paradox model in a developing-economy South Asian setting, which shows cross-cultural generalizability of the fundamental monitoring-trust-CWB mechanism besides also recording contextual differences in the magnitude of effects. Second, the joint mediation-moderation model improves on current theoretical models by demonstrating that trust is the most important mediating effect (not stress, autonomy, or justice per se) and finding gender to be a significant boundary condition. Third, the observation that digital CWB is the most common behavioral expression of monitoring reactance among

Pakistani remote workers adds to the new literature on technology-mediated CWB forms which current scales cannot adequately measure.

PRACTICAL IMPLICATIONS

The results have their direct implications on HR professionals who deal with remote workforces in Pakistan. The significant monitoring-trust relationship ($\beta = -.58$) shows that the monitoring-intensive organizations are incurring a significant trust cost that may not be compensated by the productivity gains monitoring vendors purported to have, especially with the reported indirect trust-mediated CWB costs. Companies must evaluate the monitoring portfolios of their organizations, eliminating or limiting the capabilities that they cannot explain by certain operational requirements.

The gender moderation result needs certain attention more than ever. Monitoring-related trust erosion is disproportionately experienced by female workers in Pakistani remote work settings, and organizations implementing intensive surveillance without gender-sensitive implementation practices may lose not only female talent, a rare and valuable resource in the expanding technology workforce in Pakistan, but also meet their diversity and inclusion objectives. HR managers need to take into account the gender-disaggregated monitoring impact assessment and make sure that the design processes of the monitoring policy include the meaningful representation of females in the employment.

The mediation result is that the organizational intervention that is most effective in reducing CWB in remote workforces is not increased monitoring which would further worsen the trust deficit that is a cause of CWB, but trust investment transparent communication about the purpose and limit of monitoring, participative policy design processes, and managerial leadership behaviours that convey a message of respect and organizational care to employees (Mohase et al., 2025)

LIMITATIONS

There are a number of weaknesses to this study that need to be taken into account. First, the cross-sectional design does not allow making causal inferences: although the hypothesized causal directions are theoretically well-motivated and consistent with previous longitudinal data, the current data cannot disqualify reverse causation (e.g., higher-CWB employees might be more monitored). The longitudinal or experimental research design should be used in future studies to identify the causal directionality. Second, the convenience sample, though stratified and quite large, was selected only in Lahore, Karachi, and Islamabad, and might not be representative of employees in the secondary Pakistan cities, or in other sectors other than IT and banking. Third, self-report scales of CWB are prone to social desirability bias; the current study reduced the bias by ensuring anonymity and using validated scales but future research in this area should be complemented by peer or supervisor ratings. Fourth, common method variance, in all cases where predictor and outcome variables are measured at the same time with the same respondents is of risk, was counteracted by randomizing scale and including marker variables but cannot be completely eliminated.

CONCLUSION

The aim of this study was to test the hypothesis of empirical operation of Surveillance Paradox, the self-defeating mechanism through which remote monitoring of employees kills the organizational trust that it is used to generate in the context of IT and banking industry in Pakistan. It is yes, and yes and yes. In four hypotheses where the sample size was 287 employees, organizational trust was significantly eroded by remote monitoring intensity ($\beta = -.58$), and counterproductive work behavior was significantly increased by remote monitoring intensity ($\beta = .49$), and organization trust significantly mediated the relationship between monitoring and counterproductive work behavior (indirect effect $\beta = .31$, 95% CI [.22, .4]. The

monitoring-trust relationship was moderated by gender, where female employees felt the monitoring-induced trust erosion about 42% stronger than male employees.

These results send a definite and pressing message to the Pakistani organizations that operate remote workforces: intensive monitoring is not a cheap management tool. Its trust expenses are high, its CWB-inducing effects are empirically reported, and the fact that it has a disproportionate effect on female workers raises equity as well as efficiency issues. The way ahead is not reduced responsibility but smarter responsibility - tracking that is open, proportional, participatory as well as developmental instead of clandestine, pervasive, unilateral and punitive.

The research, more generally, adds to an emerging body of international research that the industrial-era presumption which the EEM relies on, that employees need to be observed to work, is not only ethically questionable, but also empirically counterproductive. Companies that opt to use trust as their main way of accountability rather than surveillance do not lose performance; they increase performance. At the root of the Surveillance Paradox lies a management failure of imagination: the inability to perceive that sustainable organizational performance in the knowledge economy is not about the visibility of labor but the energy of the relations in which it takes place.

REFERENCES

- Alge, B. J. (2001). Effects of computer surveillance on perceptions of privacy and procedural justice. *Journal of Applied Psychology*, 86(4), 797–804. <https://doi.org/10.1037/0021-9010.86.4.797>
- Bahria University Research Group. (2025). Emotional intelligence, organizational justice, and counterproductive work behavior among Pakistani university employees. *Pakistan Journal of Humanities and Social Sciences*, 13(2), 1–18. <https://doi.org/10.52131/pjhss>
- Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85(3), 349–360. <https://doi.org/10.1037/0021-9010.85.3.349>
- Blau, P. M. (1964). *Exchange and power in social life*. John Wiley & Sons.
- Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. Academic Press.
- Brinson, N., Amini, B., Lemon, L., & Bawole, C. (2024). Electronic performance monitoring: The role of reactance, trust, and privacy concerns in predicting job satisfaction in the post-pandemic workplace. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 18(5), Article 10. <https://doi.org/10.5817/CP2024-5-10>
- Cambridge Core. (2025). The effects of being under watch: The impact of electronic monitoring on remote workers' psychological safety. *Journal of Management & Organization*. <https://doi.org/10.1017/jmo.2025.001>
- Creswell, J. W., & Creswell, J. D. (2024). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
- Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management*, 31(6), 874–900. <https://doi.org/10.1177/0149206305279602>

- Cummings, L. L., & Bromiley, P. (1996). The Organizational Trust Inventory (OTI): Development and validation. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 302–330). SAGE Publications.
- Digital Panopticon Framework. (2026). Digital panopticon: How remote work monitoring shapes employee behavior and motivation. *MDPI Behavioral Sciences*. <https://doi.org/10.3390/journal/6/1/6>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Harvard Business Review. (2024, February). Surveilling employees erodes trust — and puts managers in a bind. *Harvard Business Review*. <https://hbr.org/2024/02/surveilling-employees-erodes-trust-and-puts-managers-in-a-bind>
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). Guilford Press.
- High5Test. (2025). Employee monitoring statistics in the U.S. (2024–2025): Surveillance and AI tracking. <https://high5test.com/employee-monitoring-statistics/>
- Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44(3), 513–524. <https://doi.org/10.1037/0003-066X.44.3.513>
- Jansen, G., Noorda, H., Akkerman, A., & Manevska, K. (2026). Watched but not seen: Electronic employee monitoring as workplace imprisonment and its effect on social recognition. *Work, Employment and Society*. <https://doi.org/10.1177/00380385251359077>
- Khan, M. A., & Tariq, H. (2024). Validation of the Organizational Trust Inventory in Pakistani professional contexts. *Journal of Managerial Sciences*, 18(1), 44–61.
- Laksanadjaja, O. (2024). Uncovering the web of secrets surrounding employee monitoring software: A content analysis of information provided by vendors. *Human Behavior and Emerging Technologies*, 2024, Article 7951911. <https://doi.org/10.1155/2024/7951911>
- Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, 86(1), 114–121. <https://doi.org/10.1037/0021-9010.86.1.114>
- Marstand, A. F. (2025). 'Distant but close': Leadership behaviours, psychological distance, employee coping and effectiveness in remote work contexts. *Journal of Occupational and Organizational Psychology*, 98, e12544. <https://doi.org/10.1111/joop.12544>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>
- Mehmood, K., Jabeen, F., Iftikhar, Y., Yan, M., Khan, A. N., AlNahyan, M. T., & Alkindi, H. A. (2022). Elucidating the effects of organisational practices on innovative work behavior in UAE public sector organisations: The mediating role of employees' psychological capital. *Journal of*

- Organizational Change Management, 35(2), 286–302. <https://doi.org/10.1108/JOCM-05-2021-0132>
- Mohase, K., Donald, F., & Israel, N. (2025). Inclusive leadership, psychological safety, and employee voice in remote and hybrid work employees. *South African Journal of Psychology*. <https://doi.org/10.1177/00812463251365484>
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119–158.
- Pakistan Software Export Board. (2024). Pakistan IT industry survey and export performance report 2023–24. PSEB Annual Report. Ministry of IT & Telecom, Government of Pakistan.
- Psychological Safety in the Changing Workplace. (2024). 2024 Work in America Survey. American Psychological Association. <https://www.apa.org/pubs/reports/work-in-america/2024/psychological-safety>
- Røberg, M., Ingelsrud, M. H., & colleagues. (2025). Electronic monitoring, trust, and turnover intention in the public and private sector. *Public Management Review*. <https://doi.org/10.1080/14719037.2025.2608281>
- Searle, R. H., & Rice, C. (2025). Trust, and high control: An exploratory study of counterproductive work behaviour in a high security organization. *European Journal of Work and Organizational Psychology*, 34(3), 392–402. <https://doi.org/10.1080/1359432X.2024.2344870>
- Siegel, R., König, C. J., & Lazar, V. (2022). The impact of electronic monitoring on employees' job satisfaction, stress, performance, and counterproductive work behavior: A meta-analysis. *Computers in Human Behavior Reports*, 8, Article 100227. <https://doi.org/10.1016/j.chbr.2022.100227>
- Syed, J., & Özbilgin, M. (2019). *Managing diversity and inclusion: An international perspective*. SAGE Publications.
- Thibaut, J., & Walker, L. (1975). *Procedural justice: A psychological analysis*. L. Erlbaum Associates.
- Yang, R. (2025). Frontier exploration and future prospects of international and comparative research in human resource management. *Frontiers in Business, Economics and Management*, 21(1), 155–161. <https://doi.org/10.54097/h6r1ws15>
- Deloitte Insights. (2025). Getting to the source of the workplace trust deficit. Deloitte. <https://www.deloitte.com/us/en/insights/topics/leadership/workplace-monitoring-and-the-lack-of-trust-in-the-workplace.html>
- Gallup. (2025). The remote work paradox: Higher engagement, lower wellbeing. Gallup Workplace Research. <https://www.gallup.com/workplace/660236/remote-work-paradox-engaged-distressed.aspx>
- Pandey, S., & Mahesh, J. (2025). Emerging trends in people-centric human resource management: A systematic literature review. *Vision: The Journal of Business Perspective*. <https://doi.org/10.1177/09722629231182853>