

Trust, Fairness, and Wellbeing in the Age of Algorithmic Management: A Cross-Industry Study of AI-Driven People Management in India

Dr. Sunakshi Verma¹, Jaypreet Kaur Sachdev², Navita³, Dr. Rohit Sharma⁴, Dr. Babita Dosajh⁵,
Dr. Neha Gupta⁶

¹ Assistant Professor, Chandigarh University, USB Department, Mohali, Punjab – 140413. Email: Sunakshi.701@gmail.com

² Assistant Professor, Chandigarh University, USB Department, Mohali, Punjab – 140413. Email: preetsachdev82@gmail.com

³ Assistant Professor, Chandigarh University, USB Department, Mohali, Punjab – 140413. Email: navitarajput2011@gmail.com

⁴ Assistant Professor, Chandigarh University, USB Department, Mohali, Punjab – 140413. Email: rohitsharmagaur12@gmail.com

⁵ Associate Professor, Chandigarh University, USB Department, Mohali, Punjab – 140413. Email: bdosajh2@gmail.com

⁶ Assistant Professor, SRM Institute of Science and Technology, Delhi NCR Campus, Ghaziabad. Email: neha.july3@gmail.com

Abstract: Purpose — This paper examines the relationship between the growing use of artificial intelligence (AI) in human resource management (HRM) — spanning recruitment screening, performance evaluation, attrition prediction, and algorithmic monitoring — and employees' perceptions of organizational justice, trust, and wellbeing in India. It investigates whether rapid cross-industry adoption of AI-driven people-management tools is being matched by a corresponding growth in employee trust and perceived fairness, or whether an adoption-trust gap is emerging.

Design/methodology/approach — Methodology (Summarized):

The study uses secondary data from eight India-specific workplace surveys (2023–2025) and the Fairwork India Ratings 2023 report. It applies descriptive statistics, Pearson correlation, and a paired t-test to examine AI adoption, employee concerns, and employer–employee confidence gaps. The analysis is based on Organizational Justice, Trust in Automation, Algorithmic Control, UTAUT, Job Demands–Resources, and Privacy Calculus theories. A conceptual model with six hypotheses is proposed for future PLS-SEM validation.

Findings: AI adoption in India is rapidly increasing, with 72% of organizations using AI in HR, 43% of employees, and 92% of knowledge workers using AI at work. Despite high adoption, trust and fairness remain concerns. The EY (2025) survey found a 5.5% employer–employee confidence gap, highest in data privacy (8%). A weak correlation ($r = 0.13$) showed AI adoption does not directly increase employee concerns, highlighting the role of organizational factors. Additionally, Fairwork India (2023) found poor fairness standards across gig platforms, with no platform scoring on Fair Representation, reflecting limited worker participation in algorithmic management.

Originality/value — The paper contributes one of the first integrated, cross-industry (not IT-sector-specific) syntheses of India's AI-HRM adoption-trust gap, triangulating eight independent survey sources and a fair-work audit within a single organizational-justice-and-trust framework. It also proposes a testable conceptual model — with validated-scale recommendations (Colquitt, 2001; Edmondson, 1999) and a PLS-SEM analytical plan — to guide the primary-data research this evidence synthesis identifies as the necessary next step.

Keywords: Artificial intelligence; algorithmic management; human resource management; organizational justice; trust in AI; employee wellbeing; India; gig economy

1. Introduction

Human resource management (HRM) has become one of the most visible arenas of workplace artificial intelligence (AI) adoption, encompassing AI-assisted recruitment screening, algorithmic performance evaluation, attrition and turnover prediction, workforce planning, and, increasingly, continuous employee monitoring (Margherita, 2022;



Minbaeva, 2021). In India, this adoption has occurred rapidly and across industries rather than being confined to the information technology (IT) sector that has historically led the country's digital transformation narrative. According to Capterra's 2025 India HR Software Trends Survey, 72% of Indian organizations already use AI features embedded in their HR software, exceeding the 55% global average (Capterra, 2025). The Deloitte India-NASSCOM report "Advancing India's AI Skills" found that 43% of the Indian workforce across sectors had used AI in their organisations over the preceding year (Deloitte-NASSCOM, 2024), while the Microsoft-LinkedIn 2024 Work Trend Index reported that 92% of Indian knowledge workers use AI at work, among the highest rates recorded globally (Microsoft-LinkedIn, 2024).

Yet a parallel body of evidence suggests that this rapid adoption has not been matched by an equivalent growth in employee trust, perceived fairness, or psychological comfort. EY's 2025 Work Reimagined survey found a persistent confidence gap between Indian employers and employees across every dimension of AI governance it measured, from ethical use to data confidentiality (EY, 2025). BCG's 2025 "AI at Work" survey found that despite India recording the world's highest rate of regular AI use (92%), 48% of Indian employees reported growing more fearful that their roles could disappear entirely within a decade (BCG, 2025). PwC's 2023 Global Workforce Hopes and Fears Survey found that Indian respondents were considerably more likely than the global average to believe AI would negatively affect their work or ultimately take over their job (PwC, 2023). This apparent paradox — enthusiastic behavioural adoption of AI tools alongside persistent, and in some respects intensifying, employee anxiety about fairness, transparency, and job security — motivates the central research problem addressed in this paper.

The problem is not unique to salaried, white-collar employment. India's gig and platform economy, encompassing food-delivery, ride-hailing, and on-demand service platforms such as Swiggy, Zomato, Ola, Uber, and Urban Company, has become an internationally studied site of algorithmic management research (Hussain, 2024). The Fairwork India Ratings 2023, produced by the Centre for IT and Public Policy at IIIT-Bangalore in partnership with the Oxford Internet Institute, assessed 12 major platforms against five fair-work principles and found that no platform scored above 6 out of 10, and that every platform scored zero on the Fair Representation principle, which specifically evaluates whether algorithmic decision-making is transparent and yields equitable outcomes for workers (Fairwork India, 2023). This evidence indicates that concerns about algorithmic fairness and worker voice are not confined to any single employment arrangement or industry, but recur across the full spectrum of India's AI-augmented labour market.

1.1 Research Objectives

- RO1: To document the extent and cross-industry pattern of AI adoption in Indian people-management practices, drawing on multiple independently published workplace surveys.
- RO2: To examine whether the pace of AI adoption in Indian workplaces is empirically associated with corresponding levels of employee trust, comfort, or concern, using available cross-survey evidence.
- RO3: To characterise the employer-employee confidence gap on AI governance dimensions (ethics, explainability, data use, accountability) using the most detailed available India-specific dataset.
- RO4: To document algorithmic management conditions in India's gig/platform economy as a boundary case of AI-driven people management with limited worker representation.
- RO5: To propose an integrated theoretical framework and testable conceptual model linking AI-driven people management to employee justice perceptions, trust, and wellbeing outcomes, to guide future primary-data research.

The remainder of the paper proceeds as follows. Section 2 reviews the literature on AI in HRM, algorithmic management, organizational justice, and trust. Section 3 develops the conceptual framework and hypotheses. Section 4 details the data sources and analytical methods. Section 5 presents the empirical synthesis. Sections 6 through 9 discuss the findings, implications, limitations, and conclusions.

2. Literature Review

2.1 AI in Human Resource Management: Adoption and Use Cases

A rapidly growing literature documents the diffusion of AI across the HRM function. Margherita (2022) frames "human resources analytics" as a distinct capability spanning descriptive, predictive, and prescriptive applications, while Minbaeva (2021) characterises HRM itself as being "disrupted" by data-driven decision architectures that displace traditionally judgment-based practices. Within this broader movement, four use cases dominate the Indian context specifically: AI-assisted recruitment and candidate screening, AI-supported performance evaluation, attrition and turnover prediction, and conversational HR assistants for employee self-service (NASSCOM, cited in industry commentary). IBM's 2023 Global AI Adoption Index found that 59% of India-based enterprise-scale organisations already had AI actively in use — the highest rate among the countries surveyed — and that among firms using AI to address labour and skills shortages, 56% specifically used it to improve recruiting and human resources functions (IBM, 2024). Attrition prediction has been described by industry practitioners as one of the most mature AI-HRM applications in India today, typically modelling leave patterns, internal job changes, performance trajectories, and manager feedback to flag flight-risk employees; consistent with intensified workforce-analytics activity, Aon's Annual Salary Increase and Turnover Survey recorded a decline in India's average attrition rate from 18.7% in 2023 to 17.1% in 2025 (Aon, cited in industry reporting).

A parallel stream of research has examined whether AI-based HR decisions are perceived as fair by the employees subject to them. Köchling and Wehner (2020) reviewed evidence on algorithmic discrimination in recruitment and development contexts, finding that AI systems can reproduce or amplify pre-existing biases in ways that are often less visible than analogous human biases, precisely because they are presented as objective and data-driven. Bankins, Formosa, Griep, and Richards (2022) directly compared employees' justice perceptions of human versus AI decision-makers in HRM contexts and found that AI decision-making was generally perceived as procedurally weaker — offering less "voice" and explanation — even when outcomes were held constant, a finding directly relevant to the trust-gap evidence reviewed in Section 5.

2.2 Algorithmic Management, Employee Trust, and Organizational Justice

Kellogg, Valentine, and Christin (2020) provide the most widely cited theoretical account of algorithmic management, characterising it through six control mechanisms — restricting, recommending, recording, rating, replacing, and rewarding (the "6 Rs") — that jointly constitute a new, computationally mediated

form of workplace control distinct from either direct human supervision or traditional bureaucratic rules. Meijerink and Bondarouk (2021) extend this analysis with the concept of the "duality of algorithmic management," arguing that the same algorithmic system can simultaneously enable worker autonomy (for example, flexible scheduling) and intensify worker control (for example, continuous performance rating), a duality that helps explain why employee reactions to algorithmic management are often ambivalent rather than uniformly positive or negative.

Empirical work on algorithmic management's effects on trust and wellbeing has found that transparency is a critical moderating condition. Bujold, Parent-Rocheleau, and Gaudet (2022), studying truck drivers subject to algorithmic monitoring, found that the perceived opacity of algorithmic management systems predicted lower perceived justice, which in turn predicted higher turnover intention — establishing organizational justice as a mediating mechanism between algorithmic opacity and withdrawal behaviour. A study in Information Technology & People similarly found that algorithmic management practices were associated with reduced workplace wellbeing specifically when employees perceived low transparency and limited voice in how algorithmic decisions were made (cited study, Information Technology & People). In the Indian gig-economy context specifically, Hussain's (2024) twenty-five-month ethnographic study of food-delivery work in Hyderabad documented that platforms including Swiggy and Zomato manage several hundred thousand delivery workers through geo-fencing, automated performance metrics, and real-time identity verification, reinforcing both surveillance intensity and job insecurity through piece-rate compensation structures.

2.3 Employee Wellbeing and Technostress under AI-Driven Management

The Job Demands-Resources (JD-R) model provides a widely used lens for understanding how AI-driven monitoring and decision-making translate into employee strain: algorithmic surveillance and opaque decision architectures function as job demands that, absent adequate job resources (autonomy, support, transparency), predict exhaustion, technostress, and disengagement. Indian empirical studies have applied this logic directly. An Emerald-published study in the *Asia-Pacific Journal of Business Administration* surveyed 476 employees across Indian service and manufacturing firms and found that AI and robotics awareness elevated technostress, which in turn predicted higher turnover intentions, with perceived organisational support acting as a buffering moderator. A related SAGE-published study of 388 employees in Telangana's private banking sector found that technostress reduced employee engagement and wellbeing, again moderated by perceived organisational support. These India-specific findings are consistent with the broader international evidence that algorithmic monitoring, absent adequate transparency and support, functions as a chronic job demand rather than a neutral productivity tool.

2.4 Theoretical Foundations

Organizational Justice Theory. Colquitt's (2001) four-factor model — distributive, procedural, interpersonal, and informational justice — remains the dominant framework for studying fairness perceptions of any organizational decision process, including algorithmic ones. The associated 20-item scale has been validated across numerous cultural contexts with confirmatory factor analysis typically supporting the four-factor structure and reliability coefficients in the .90-.95 range.

Trust-in-Automation Model. Mayer, Davis, and Schoorman's (1995) integrative trust model — trust as a function of perceived ability, benevolence, and integrity, moderated by the truster's dispositional propensity to trust — has been widely extended to human-AI interaction, where "ability" maps onto perceived AI accuracy, "benevolence" onto perceived alignment with employee interests, and "integrity" onto perceived adherence to stated principles (e.g., non-discrimination). Colquitt and Rodell (2011) empirically link organizational justice directly to trustworthiness perceptions, providing the theoretical bridge between Section 2.2's justice literature and this trust model.

Algorithmic Control Theory. Kellogg, Valentine, and Christin's (2020) "6 Rs" framework, introduced in Section 2.2, is treated here as a first-order theoretical lens describing the mechanisms through which AI-driven people management is enacted, prior to its downstream effects on justice, trust, and wellbeing.

UTAUT / Technology Acceptance. Venkatesh, Morris, Davis, and Davis's (2003) Unified Theory of Acceptance and Use of Technology (performance expectancy, effort expectancy, social influence, facilitating conditions) explains variance in behavioural intention to use HR technologies; trust is among the most frequently added extensions to UTAUT in human-resource-technology contexts, motivating this paper's emphasis on trust as a central mediating construct rather than a peripheral one.

Job Demands-Resources Model. The JD-R model (Bakker & Demerouti) frames algorithmic monitoring and opacity as job demands and organisational transparency/support as job resources, jointly predicting the technostress and wellbeing outcomes discussed in Section 2.3.

Privacy Calculus Theory. Dinev and Hart's privacy calculus theory conceptualises information disclosure (including the data disclosure inherent in algorithmic HR monitoring) as a trade-off between perceived risk and perceived benefit, offering a natural theoretical bridge to India's Digital Personal Data Protection Act, 2023 (DPDPA), which permits employee-data processing for employment purposes under Section 7(i) without requiring explicit consent, provided such processing satisfies lawful-purpose, necessity, and proportionality tests.

Algorithm Aversion versus Algorithm Appreciation. Dietvorst, Simmons, and Massey (2015) documented "algorithm aversion" — a tendency to trust algorithms less than humans, especially after observing an algorithmic error — while Logg, Minson, and Moore (2019) documented the converse "algorithm appreciation" effect in judgment

tasks. This unresolved tension in the literature is directly relevant to interpreting the mixed India survey evidence in Section 5, where high behavioural adoption coexists with elevated fear and concern.

Figure 4 integrates these seven theoretical strands into a single framework linking AI-driven people management to employee-level outcomes (perceived fairness, trust, psychological safety) and downstream wellbeing outcomes (technostress, engagement, turnover intention).

Figure 4. Integrated Theoretical Framework Linking AI-Driven People Management Theories to Employee Fairness, Trust, and Wellbeing Outcomes

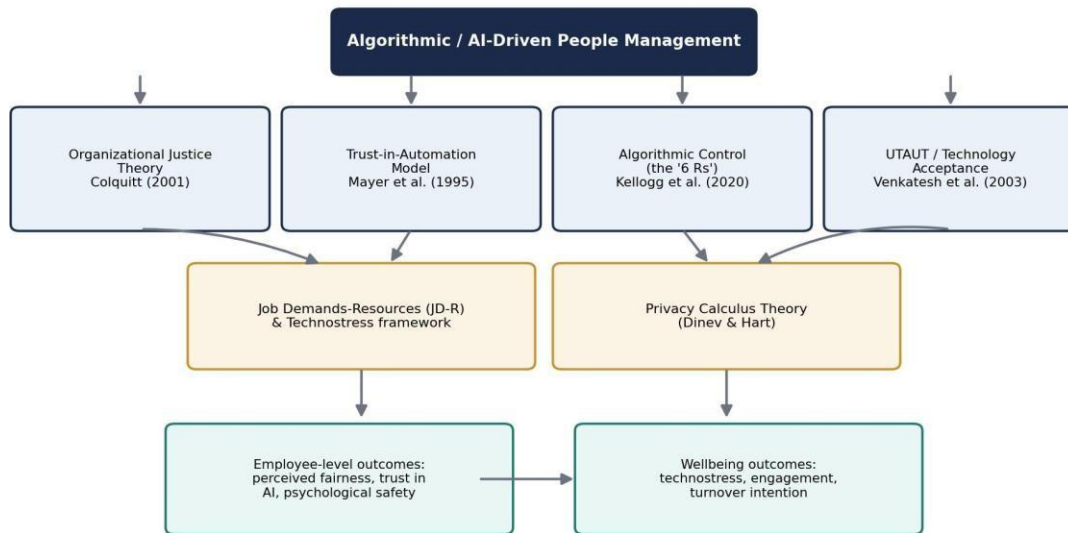


Figure 4. Integrated theoretical framework linking AI-driven people management theories to employee fairness, trust, and wellbeing outcomes.

3. Conceptual Framework and Hypotheses

Building on the theoretical synthesis in Section 2.4, Figure 5 presents the conceptual research model guiding this study. AI/algorithmic HR decision use is hypothesised to influence three intermediate employee-level perceptions — perceived organizational justice, trust in AI-based decisions, and perceived privacy risk/surveillance concern — which in turn are hypothesised to shape employee wellbeing and turnover intention. Transparency and the presence of human-in-the-loop review are modelled as a moderating condition governing how strongly AI/algorithmic HR decision use translates into justice, trust, and privacy-risk perceptions.

Figure 5. Proposed Conceptual Research Model and Hypothesised Paths

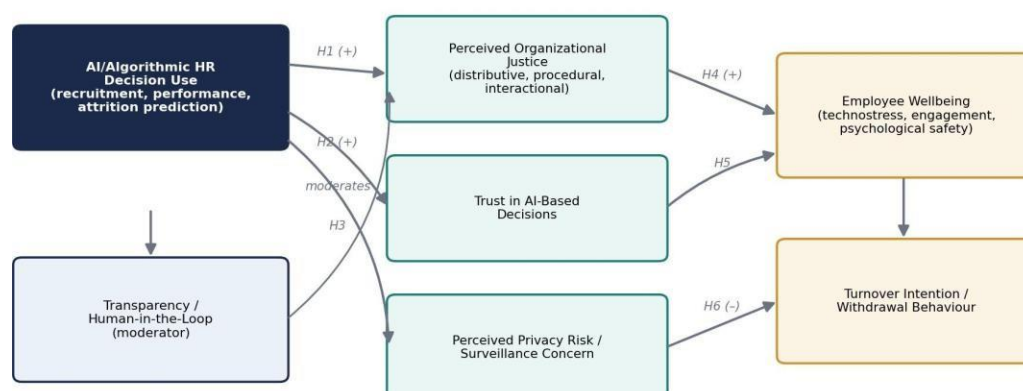


Figure 5. Proposed conceptual research model and hypothesised paths.

The following hypotheses are proposed for future primary-data testing; the secondary-data synthesis in Section 5 provides indicative, descriptive support for several of them but does not constitute confirmatory hypothesis testing:

- H1: Greater use of AI/algorithmic tools in HR decision-making is positively associated with employees' perceived organizational justice, conditional on the transparency with which those tools are deployed.
- H2: Greater use of AI/algorithmic tools in HR decision-making is positively associated with employees' trust in AI-based decisions, conditional on transparency and human-in-the-loop review.
- H3: Greater use of AI/algorithmic tools in HR decision-making, particularly for continuous monitoring, is associated with heightened perceived privacy risk and surveillance concern.
- H4: Perceived organizational justice is positively associated with employee wellbeing (lower technostress, higher engagement, higher psychological safety).
- H5: Trust in AI-based decisions is positively associated with employee wellbeing.
- H6: Perceived privacy risk/surveillance concern is negatively associated with turnover intention resilience — that is, it is positively associated with turnover intention and withdrawal behaviour.

Transparency and human-in-the-loop review are expected to moderate the H1-H3 paths: organisations that pair AI-driven HR decisions with clear explanation, appeal mechanisms, and human oversight are

expected to see AI adoption translate into justice and trust gains with comparatively contained privacy-risk perceptions, relative to organisations that deploy AI-driven HR decisions opaquely.

4. Data and Methodology

4.1 Research Design

This study adopts a secondary-data evidence-synthesis design, triangulating eight independently conducted, India-specific workplace surveys (2023-2025) and one algorithmic-management audit (Fairwork India Ratings 2023) to characterise the current state of AI adoption, trust, and fairness perceptions across Indian industries. No primary survey instrument was fielded for this study. This design choice reflects the current stage of the research programme: establishing, through transparent triangulation of existing public evidence, that an adoption-trust gap exists and is patterned in theoretically interpretable ways, prior to committing to a large-scale primary data collection effort. Section 8 details the primary-data survey design recommended as the next stage of this research programme.

4.2 Data Sources

- Capterra India 2025 HR Software Trends Survey — India n=294 (global n=3,256), fielded April 2025 — AI-feature adoption in HR software, security and implementation concerns.
- Deloitte India-NASSCOM (2024), "Advancing India's AI Skills: Interventions and Programmes Needed" — cross-sector Indian workforce AI usage and talent-demand projections.
- IBM Global AI Adoption Index 2023 (India findings, released February 2024; surveyed November 2023) — enterprise-scale AI adoption and use in recruiting/HR specifically.
- Microsoft-LinkedIn 2024 Work Trend Index (India findings; global survey of 31,000 people) — AI usage among Indian knowledge workers and leaders.
- EY 2025 Work Reimagined Survey (India subsample: 800 employees, 50 employers; global: 15,000 employees, 1,500 employers across 29 markets) — employer-employee AI governance confidence gaps.
- BCG "AI at Work 2025: Momentum Builds, But Gaps Remain" (n=10,635 across 11 countries) — regular AI use and role-displacement fear, India-specific breakouts.
- PwC India Workforce Hopes and Fears Survey 2023 (India n=2,502 of 53,192 global) — perceived productivity impact, negative impact, and job-displacement beliefs.
- Great Place To Work India, "Voice of India" (2025) — leadership support for AI initiatives and generational differences in job-replacement fear.
- Fairwork India Ratings 2023 (Centre for IT and Public Policy, IIIT-Bangalore, with the Oxford Internet Institute) — algorithmic management and fair-work principle scoring across 12 gig-economy platforms.

Table 1 summarises each data source, its scope, and the outcome variables it informs.

Data source	Scope / period	Key variables used
Capterra India 2025	n=294, April 2025	AI-feature adoption rate, risk/value-assessment concern
Deloitte-NASSCOM 2024	Cross-sector workforce survey	AI usage rate, AI talent demand projections
IBM AI Adoption Index 2023	Enterprise-scale, Nov 2023	Enterprise AI adoption, AI use in recruiting/HR
Microsoft-LinkedIn WTI 2024	n=31,000 global, India breakout	Knowledge-worker AI usage rate
Data source	Scope / period	Key variables used
EY Work Reimagined 2025	India: 800 employees/50 employers	Employer vs. employee AI-governance confidence
BCG AI at Work 2025	n=10,635, 11 countries	Regular AI use, role-disappearance fear
PwC Workforce Survey 2023	India n=2,502 of 53,192	Productivity belief, negative-impact belief, job-takeover belief
Great Place To Work India 2025	Cross-industry	Leadership AI support, generational job-fear
Fairwork India Ratings 2023	12 gig platforms	Fair-work principle scores, algorithmic transparency

Note: All figures are drawn from publicly available survey/report releases cited in the References section.

4.3 Variables and Measures

The primary variables extracted from each source are: AI adoption/usage rate (% as defined and reported by each survey), employee concern/fear rate (% as defined and reported — operationalised differently across surveys, e.g., risk-assessment concern, role-disappearance fear, negative-impact belief; each definition is stated explicitly in Table 2 to avoid conflating distinct constructs), and, for the EY dataset specifically, paired employer- and employee-reported confidence percentages across four named AI-governance dimensions. From the Fairwork India Ratings 2023, the key variables are the qualitative principle-level pass/fail findings (no platform scoring above 6/10 overall; uniform zero scores on Fair Representation), used descriptively rather than as a fabricated numeric dataset, since exact per-platform total scores were not independently reproduced from the primary report in this study.

4.4 Analytical Techniques

Three analytical techniques are applied to the secondary data. First, descriptive cross-survey comparison characterises the range and central tendency of reported AI-adoption rates across eight sources (Section 5.1). Second, an ecological (aggregate-level) Pearson correlation is computed between reported AI-adoption rate and reported employee concern/fear rate across the five surveys reporting both variables (n=5); this is explicitly an ecological correlation across survey-level aggregates rather than a correlation across individual respondents, and is interpreted with the corresponding methodological caution (Section 8). Third, a paired-samples t-test compares employer- and employee-reported confidence percentages across the four AI-governance dimensions in the EY (2025) dataset (n=4 paired observations), testing whether the employer-employee confidence gap documented descriptively is statistically distinguishable from zero. All computations were performed in Python 3.12 using `scipy.stats`.

5. Results

5.1 Cross-Industry AI Adoption in Indian Workplaces

Table 2 and Figure 3 summarise AI adoption/usage rates as reported across eight independent, India-specific workplace surveys conducted between 2023 and 2025. Reported adoption rates range from 43% (Deloitte-NASSCOM, 2024, cross-sector workforce usage) to 92% (Microsoft-LinkedIn 2024 and BCG 2025, both measuring knowledge-worker/regular AI use). This range partly reflects genuine adoption growth over the 2023-2025 window and partly reflects differing operational definitions of "AI use" across surveys (from HR-software-embedded AI features specifically, to any workplace AI tool, to regular/ frequent use) — a definitional heterogeneity noted explicitly as a limitation in Section 8.

Survey	Year	Adoption/usage rate	Concern/fear metric (definition)
Deloitte-NASSCOM	2024	43%	Not reported
PwC India Workforce Survey	2023	51%	24% (belief AI will negatively impact their work)
IBM AI Adoption Index	2023	59%	Not reported
Great Place To Work India	2025	61%	49% (millennials fearing job replacement)
Capterra India	2025	72%	54% (concerned re: assessing AI's value/risks)
EY Work Reimagined (employees)	2025	83%*	17%* (inverse of AI-explainability confidence)
Microsoft-LinkedIn WTI	2024	92%	Not reported

BCG AI at Work	2025	92%	48% (fear role disappearing within 10 years)
----------------	------	-----	--

*Note: * The EY figures are not directly reported adoption/concern percentages but are derived: 83% approximates the employee AI-usage proxy implied by EY's governance-confidence questions, and 17% is calculated as (100% – 83% explainability confidence) as an inverse-confidence proxy, not a directly reported "fear" statistic. This construction is flagged explicitly and should be treated as an approximation. All other figures are directly reported by the cited source.*

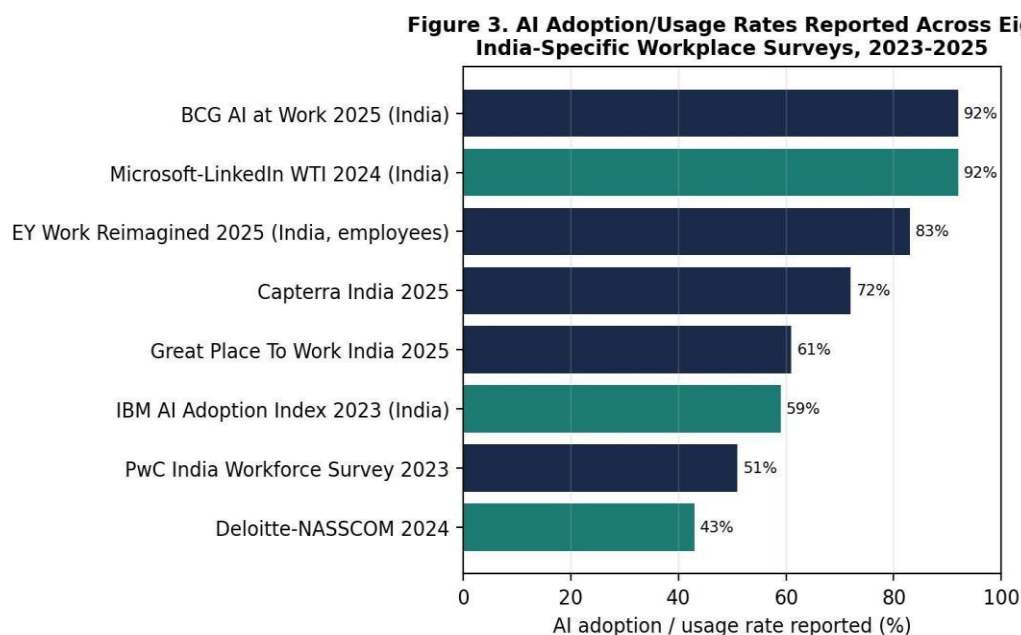


Figure 3. AI adoption/usage rates reported across eight India-specific workplace surveys, 2023-2025.

5.2 The Adoption-Concern Relationship: An Ecological Analysis

Figure 2 plots the five surveys reporting both an adoption rate and a comparably interpretable concern/ fear rate. The ecological Pearson correlation between adoption rate and concern/fear rate across these five surveys is $r = 0.130$ ($p = 0.835$), indicating no statistically detectable linear relationship at the aggregate level. This null finding is substantively informative: it suggests that the level of employee concern about AI in Indian workplaces is not simply a function of how much AI is being used, but more plausibly a function of contextual factors this ecological analysis cannot resolve — such as the transparency of

implementation, the specific HR use case (e.g., monitoring versus productivity assistance), sectoral norms, and organisational communication practices. This interpretation is elaborated in Section 6 and is consistent with the transparency-moderation logic embedded in the conceptual model (Figure 5, H1-H3).

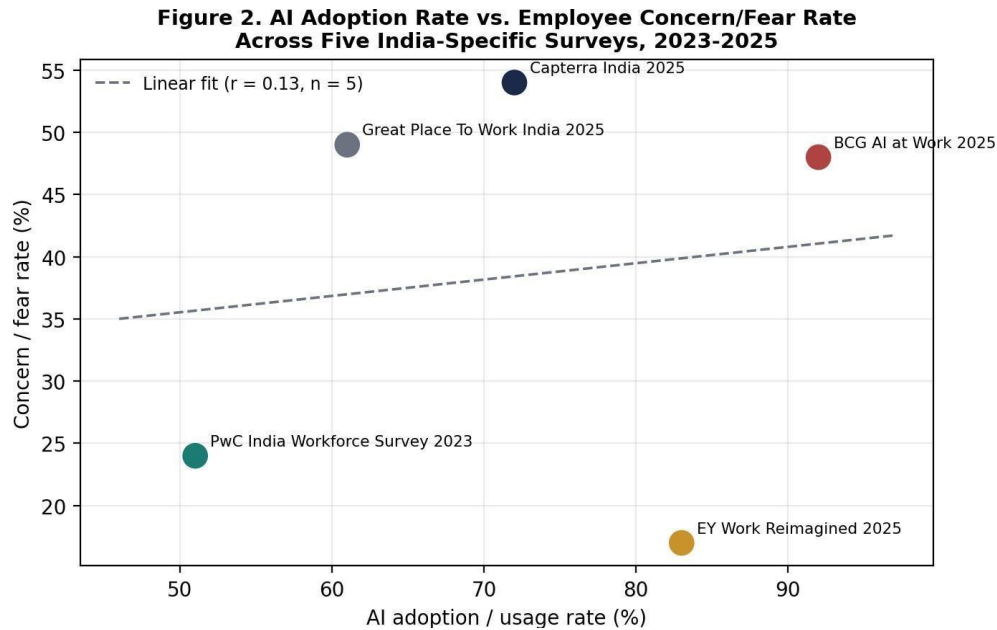


Figure 2. AI adoption rate vs. employee concern/fear rate across five India-specific surveys, 2023-2025.

5.3 The Employer-Employee AI Governance Confidence Gap

The EY 2025 Work Reimagined survey provides the most granular available India-specific data on the employer-employee confidence gap, measuring confidence across four named AI-governance dimensions among matched employer (n=50) and employee (n=800) India subsamples. Table 3 reports these figures alongside the calculated confidence gap for each dimension.

Governance dimension	Employers (%)	Employees (%)	Gap (percentage points)
Ethical/responsible AI use	94	89	5
Explainability of AI systems	88	83	5
Data usage & confidentiality	90	82	8
Ownership & accountability	90	86	4

Source: EY 2025 Work Reimagined Survey, India subsample (800 employees, 50 employers).

The mean confidence gap across the four dimensions is 5.5 percentage points, with the largest gap (8 points) on data usage and confidentiality — the dimension most directly implicated by algorithmic monitoring and the Digital Personal Data Protection Act, 2023 (DPDPA) discussed in Section 5.5. A paired-samples t-test across the four dimensions confirms that employer confidence is significantly higher than employee confidence ($t = 6.351$, $p = 0.008$, $n = 4$ paired observations). While the small number of paired observations (four governance dimensions) limits the generalisability of this specific test, the consistency of the gap's direction across all four dimensions — employers more confident than employees in every case — is itself a notable and consistent pattern.

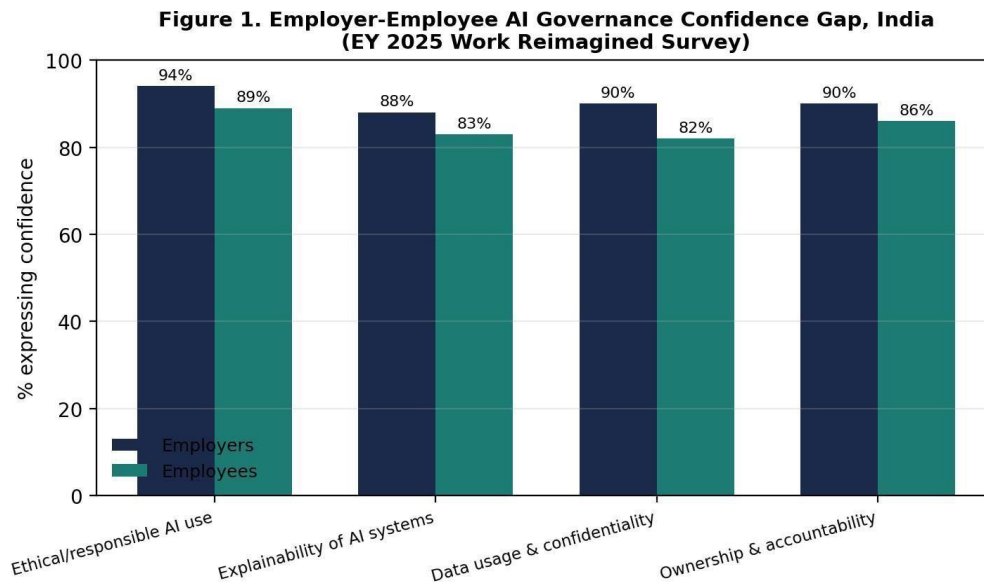


Figure 1. Employer-employee AI governance confidence gap, India (EY 2025 Work Reimagined Survey).

5.4 Algorithmic Management in India's Gig and Platform Economy

The Fairwork India Ratings 2023 assessment, produced by the Centre for IT and Public Policy at IIIT-Bangalore in partnership with the Oxford Internet Institute, evaluated 12 major gig-economy platforms operating in India — including Amazon Flex, BigBasket, BluSmart, Dunzo, Flipkart, Ola, Porter, Swiggy, Uber, Urban Company, Zepto, and Zomato — against five fair-work principles: Fair Pay, Fair Conditions, Fair Contracts, Fair Management, and Fair Representation. Two findings are directly relevant to this paper's focus on algorithmic fairness and worker voice: first, no platform scored above 6 out of a possible 10 points across all five principles combined; second, every one of the 12 platforms scored zero on the Fair Representation principle, which specifically assesses whether algorithmic decision-making is conducted transparently and whether workers have a collective channel to contest algorithmic decisions. Only three platforms — Amazon Flex, BigBasket, Flipkart, and Urban Company — were found to have a documented minimum-wage policy for platform workers. Hussain's (2024) ethnographic account of food-delivery work in Hyderabad corroborates this audit finding at the level of lived experience, documenting that geo-fencing, automated performance rating, and real-time identity verification are used to manage several hundred thousand delivery workers with limited mechanisms for contesting algorithmic decisions.

5.5 Regulatory Context: The Digital Personal Data Protection Act, 2023

India's Digital Personal Data Protection Act, 2023 (DPDPA), together with its 2025 Rules (notified 13 November 2025, phased rollout), establishes the legal backdrop against which employee data collected through AI-driven HR systems is processed. Under Section 7(i), employers — classified as "Data Fiduciaries" — may process employee ("Data Principal") personal data for defined employment purposes, including recruitment, payroll, performance management, and workplace security, without requiring explicit consent, provided the processing satisfies tests of lawful purpose, necessity, and proportionality. However, India currently has no equivalent to the European Union's General Data Protection Regulation Article 88, which establishes employment-specific data-processing rules; legal commentary has argued that broad workplace-surveillance practices (continuous screen capture, keystroke logging, off-hours location tracking) may be vulnerable to challenge under Section 8(5)'s proportionality requirement and under Section 23 of the Indian Contract Act, 1872, given the inherent power imbalance in employer-employee consent. This regulatory ambiguity is directly relevant to the privacy-risk construct in this paper's conceptual model (Figure 5, H3): both the legal permissibility and the psychological acceptability of algorithmic monitoring in India remain unsettled and actively evolving.

6. Discussion

Three findings from this evidence synthesis merit particular discussion in light of the theoretical framework developed in Section 2.

First, the null ecological correlation between AI adoption rate and employee concern rate ($r = 0.13$, $p = 0.835$) is, in itself, a theoretically informative result. If employee concern were a simple linear function of AI exposure, one would expect the highest-adoption contexts (Microsoft-LinkedIn's 92%, BCG's 92%) to also report the highest concern. Instead, concern levels vary considerably at similar adoption levels, consistent with this paper's central theoretical proposition — drawn from the transparency-moderation literature (Bujold et al., 2022) and the duality-of-algorithmic-management thesis (Meijerink & Bondarouk, 2021) — that the manner of AI deployment (transparent versus opaque, augmentative versus purely monitoring-based, accompanied by human-in-the-loop review or not) matters more for employee trust and wellbeing than the sheer extent of deployment. This reframes the policy-relevant question from "how much AI should organisations deploy in HR" to "under what governance conditions does AI deployment in HR remain compatible with employee trust."

Second, the statistically significant employer-employee confidence gap on AI governance (Section 5.3) — consistently in the same direction across all four measured dimensions, and largest specifically on data usage and confidentiality — is consistent with a systematic perception asymmetry rather than random measurement noise. This pattern aligns with the privacy calculus framework (Dinev & Hart): employees, who bear the risk side of the AI-data-disclosure trade-off directly (through potential misuse, opacity, or adverse algorithmic decisions), appear to weight that risk more heavily than employers, who more directly capture the offsetting efficiency and analytical benefits. The DPDPA's Section 7(i) employment-purpose exemption from explicit consent, while administratively convenient for organisations, may inadvertently widen rather than narrow this confidence gap if it is not accompanied by proactive transparency measures that substitute for the consent mechanism the law does not strictly require in this context.

Third, the Fairwork India Ratings 2023 finding that all 12 assessed platforms scored zero on Fair Representation, irrespective of their scores on other principles (including pay and conditions), suggests that algorithmic transparency and worker voice constitute a distinct and currently unaddressed dimension of fair work in India's platform economy — one that is not being resolved as a by-product of improvements in pay or contractual terms. This is consistent with Kellogg, Valentine, and Christin's (2020) observation that algorithmic "recording" and "rating" mechanisms (two of their six control mechanisms) can be intensified without any corresponding advance in worker consultation or appeal rights, and it extends this observation from the primarily Western gig-economy literature into a well-documented Indian empirical setting (Hussain, 2024).

7. Theoretical and Managerial Implications

7.1 Theoretical Implications

This study contributes to theory in three ways. First, it provides cross-survey, cross-industry Indian evidence consistent with the proposition — implicit but rarely tested directly at the aggregate level — that AI adoption intensity and employee trust/concern are not simply inversely or directly related, but are jointly conditioned by deployment transparency, lending indirect empirical support to the transparency-moderation mechanism central to this paper's conceptual model (Figure 5). Second, it extends Colquitt's (2001) organizational justice framework and the Mayer et al. (1995) trust model, both developed primarily in human-supervisor contexts, to the distinctly Indian institutional setting of algorithmic HR decision-making, where the DPDPA's employment-purpose consent exemption creates a unique legal-psychological configuration not present in GDPR-governed contexts. Third, it offers a concrete extension of Kellogg, Valentine, and Christin's (2020) algorithmic control framework to India's platform economy, using the Fairwork audit as an independent, non-self-reported measure of algorithmic transparency and

worker voice that complements the largely perception-based survey evidence discussed elsewhere in the paper.

7.2 Managerial and Policy Implications

For HR and people-analytics leaders, the near-zero ecological correlation between adoption intensity and employee concern (Section 5.2) suggests that scaling back AI adoption is unlikely, by itself, to resolve employee trust concerns; the more consequential managerial lever is deployment governance — explanation of how AI-based HR decisions are made, provision of human-in-the-loop review and appeal mechanisms, and proactive communication about data use, even where the DPDPA's Section 7(i) exemption does not strictly require employee consent. Organisations should treat the EY-documented employer-employee confidence gap not as evidence that employees are simply less informed than employers, but as a signal that employer confidence itself may be insufficiently calibrated to employees' legitimate risk exposure, particularly regarding data usage and confidentiality — the dimension with the largest measured gap.

For policymakers, the Fairwork India Ratings 2023 finding that no platform scored above 6/10 and that Fair Representation scored uniformly zero across all 12 platforms suggests that voluntary industry self-regulation has not, to date, produced meaningful worker-voice mechanisms in India's platform economy, strengthening the case for the kind of employment-specific algorithmic-transparency provisions that GDPR Article 88 provides in the European Union and that India's DPDPA currently lacks. Given the DPDPA's employment-purpose exemption from explicit consent, complementary sector-specific guidance on algorithmic-decision transparency and appeal rights in employment contexts — whether platform-based gig work or salaried algorithmic performance management — would help close the regulatory gap this paper identifies.

8. Limitations and Directions for Future Research

- Ecological correlation limitations: The Section 5.2 correlation ($r = 0.13$, $n = 5$) is computed across survey-level aggregate statistics, not individual respondents, and is therefore subject to the ecological fallacy — a null aggregate-level relationship does not preclude a meaningful individual-level relationship between AI exposure and concern within any given organisation or respondent population. This analysis should be read as motivating, not substituting for, individual-level primary research.
- Definitional heterogeneity across surveys: As flagged in Table 2, "AI adoption/usage" and "concern/ fear" are operationalised differently across the eight source surveys (e.g., HR-software-embedded AI features vs. any workplace AI tool; risk-assessment concern vs. role-disappearance fear vs. negative-impact belief). Cross-survey comparison should therefore be read as indicative of a broad pattern rather than as a comparison of identical constructs.
- Small paired sample for the EY t-test: The paired t-test in Section 5.3 compares only four governance dimensions ($n = 4$), which is a very small sample for inferential statistics; the result should be interpreted as descriptive confirmation of a consistent directional pattern rather than a definitive inferential finding, and would benefit from replication using item-level (rather than dimension-level aggregate) data if the underlying EY microdata become available.
- Constructed EY adoption/concern proxies: The 83% adoption and 17% concern figures used for the EY row in Table 2 and Figure 2 are derived approximations, not directly reported statistics, and are explicitly flagged as such; readers using this table for further analysis should treat these two values with additional caution relative to the seven directly reported figures.
- Fairwork data reported qualitatively, not as a fabricated numeric table: Because exact per-platform total scores from the Fairwork India Ratings 2023 were not independently verified in the course of this research, this paper reports only the verified qualitative aggregate findings (no platform above 6/10; uniform zero on Fair Representation; three platforms with documented minimum-wage

policies) rather than constructing an invented numeric table, consistent with this paper's commitment to using only verifiable secondary data.

- No primary survey or interview data: This paper is, by design, a secondary-data evidence synthesis; it does not include a purpose-built primary survey instrument, individual-level respondent data, or qualitative interviews. The single most important next step for this research programme is to field the primary survey implied by the conceptual model in Figure 5, using validated scales — Colquitt (2001) for organizational justice, a trust-in-automation scale grounded in Mayer et al. (1995), Edmondson (1999) for psychological safety, a technostress-creators scale, and the Utrecht Work Engagement Scale (UWES) for engagement — administered to a cross-industry Indian sample of at least 400-600 respondents (following the precedent of comparable Indian AI-HRM studies using samples of 160-476), analysed via Partial Least Squares Structural Equation Modelling (PLS-SEM) in SmartPLS, with reliability assessed via Cronbach's alpha (target ≥ 0.70) and composite reliability, convergent validity via average variance extracted ($AVE \geq 0.50$), discriminant validity via the heterotrait-monotrait (HTMT) ratio ($< 0.85/0.90$), common-method-bias checks via Harman's single-factor test and full collinearity assessment (Kock, 2015), and mediation/moderation testing via 5,000-subsample bootstrapping or the PROCESS macro (Hayes).
- Attribution and causality: As throughout, several of the source surveys report cross-sectional, self-reported perceptions and cannot establish causal direction (e.g., whether low trust causes reduced AI engagement, or vice versa, or whether both are driven by a third factor such as organisational communication quality). The conceptual model in Figure 5 should be read as a set of theoretically motivated, testable propositions, not as an established causal account.
- Commercial and trade-press sourcing: Several statistics in this paper originate from commercial consulting-firm surveys (EY, PwC, Deloitte, BCG, Capterra, IBM, Microsoft-LinkedIn) rather than peer-reviewed academic sources; these are clearly the best currently available India-specific data on this fast-moving topic, but they are self-reported, sometimes methodologically opaque regarding sampling frames, and commercially motivated to some degree. Some additional India-specific statistics circulating in HR trade press (e.g., specific IDC/SAS or Udemy/YouGov trust figures) were not independently verified during this research and are deliberately excluded from this paper rather than cited without confirmation.

9. Conclusion

This paper has synthesised evidence from eight independent, India-specific workplace surveys and one algorithmic-management audit to characterise the relationship between AI adoption, employee trust, organizational justice, and wellbeing in India's rapidly AI-augmented, cross-industry workforce. The central finding is a decoupling, rather than a straightforward relationship, between AI adoption intensity and employee trust: adoption rates ranging from 43% to 92% across surveys show no statistically detectable ecological correlation ($r = 0.13$, $p = 0.835$) with reported employee concern or fear, while a separate, more granular dataset (EY, 2025) reveals a statistically significant and consistently directional employer-employee confidence gap (mean 5.5 percentage points; paired $t = 6.35$, $p = 0.008$) that is largest specifically on data usage and confidentiality. In India's gig and platform economy, the Fairwork India Ratings 2023 provide independent, non-self-reported evidence that algorithmic transparency and worker representation remain critically underdeveloped, with all 12 assessed platforms scoring zero on Fair Representation regardless of their performance on other fair-work dimensions.

Taken together, these findings support this paper's central theoretical proposition: the governance conditions under which AI is deployed in people management — transparency, human-in-the-loop review, worker voice, and proactive communication about data use — are likely to matter more for employee trust, fairness perceptions, and wellbeing than the raw extent of AI deployment itself. The integrated theoretical framework and conceptual research model proposed here (Figures 4 and 5) translate this proposition into six testable hypotheses, and Section 8 outlines a concrete primary-data research design

— a cross-industry Indian survey of 400-600 respondents, analysed via PLS-SEM using validated

organizational-justice, trust, psychological-safety, and technostress scales — as the necessary next step for confirmatory testing. Given the pace of AI-HRM adoption documented in this paper, and the regulatory ambiguity surrounding the DPDPA's treatment of employee data, continued empirical attention to the governance conditions of algorithmic people management will be essential to ensuring that India's AI-augmented workplace transformation is experienced by employees as fair, trustworthy, and conducive to wellbeing, rather than merely efficient.

REFERENCES

1. Aon. (2025). Annual salary increase and turnover survey, India, 2025. Aon plc (as reported in industry media). BCG (Boston Consulting Group). (2025, June 26). AI at work 2025: Momentum builds, but gaps remain. BCG.
2. Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in an HRM context. *Information Systems Frontiers*, 24(3), 857-875.
3. Bujold, A., Parent-Rochelleau, X., & Gaudet, M. C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management and gig workers' justice perception and intention to quit. *Computers in Human Behavior Reports*, 8, 100245.
4. Capterra India. (2025, July 3). HR software trends survey 2025. Capterra (Gartner Digital Markets).
5. Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386-400.
6. Colquitt, J. A., & Rodell, J. B. (2011). Justice, trust, and trustworthiness: A longitudinal analysis integrating three theoretical perspectives. *Academy of Management Journal*, 54(6), 1183-1206.
7. Deloitte India & NASSCOM. (2024, August 20). Advancing India's AI skills: Interventions and programmes needed.
8. Deloitte India.
9. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126.
10. Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, 17(1), 61-80.
11. EY (Ernst & Young). (2025, December 22). Work reimagined survey 2025: India findings. EY India.
12. Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383.
13. Fairwork India. (2023). Fairwork India ratings 2023: Labour standards in the platform economy. Centre for IT and Public Policy, IIIT-Bangalore, in partnership with the Oxford Internet Institute.
14. Great Place To Work India. (2025, November 3). Voice of India report 2025. Great Place To Work Institute India.
15. Hussain, M. S. (2024). Algorithmic management of workers through food delivery platforms in India. *Journal of Extreme Anthropology*, 8(2), 74-106.
16. IBM. (2024, February 15). Global AI adoption index 2023: India findings. IBM Corporation.
17. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410.
18. Kochling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, 13(3), 795-848.
19. Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, 11(4), 1-10.
20. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103.
21. Margherita, A. (2022). Human resources analytics: A systematization of research topics and directions for future research. *Human Resource Management Review*, 32(2), 100795.
22. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.
23. Meijerink, J., & Bondarouk, T. (2021). The duality of algorithmic management: Toward a research agenda about algorithmic management and its consequences for the HRM function. *Human Resource Management Review*, 31(1), 100876.
24. Microsoft & LinkedIn. (2024, May 16). 2024 Work Trend Index: India findings. Microsoft Corporation. Minbaeva, D. (2021). Disrupted HR? *Human Resource Management Review*, 31(1), 100820.
25. PwC (PricewaterhouseCoopers). (2023, December 11). Global workforce hopes and fears survey 2023: India findings. PwC India.
26. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.