

Responsible AI-Enabled Human Resource Management and Employee Autonomy: The Mediating Role of Trust in AI and Psychological Safety

Authors: Dishaa Vyas¹, Prastuti Gupta²

Assistant Professor; Sri Aurobindo Institute of Technology, Indore, MP^{1,2}
dishaa.vyas@sait.ac.in¹, prastuti.gupta@sait.ac.in²

ABSTRACT

Artificial intelligence (AI) is increasingly transforming human resource management (HRM) through algorithm-supported recruitment, performance management, employee monitoring, workforce planning, and employee development. While existing research has examined the efficiency, fairness, transparency, and ethical implications of AI-enabled HRM, comparatively limited attention has been given to how such systems influence employees' willingness to express suggestions, concerns, and problems. This study examines the relationship between AI-enabled human resource management and employee voice by investigating psychological safety as a mediating mechanism and perceived human oversight as a moderating condition. A quantitative, cross-sectional research design was adopted, and data were obtained from the employees working in Indian organisations using AI-enabled or algorithm-supported HR practices. The proposed relationships were examined using Partial Least Squares Structural Equation Modelling (PLS-SEM) with bootstrapping. The results indicate that AI-enabled HRM is positively associated with psychological safety ($\beta = 0.421$, $p < 0.001$), while psychological safety has a significant positive relationship with employee voice ($\beta = 0.486$, $p < 0.001$). AI-enabled HRM also demonstrates a significant direct relationship with employee voice ($\beta = 0.218$, $p < 0.001$), as well as with promotive and prohibitive voice. Psychological safety significantly mediates the relationship between AI-enabled HRM and employee voice, including both promotive and prohibitive voice. Perceived human oversight significantly strengthens the relationship between AI-enabled HRM and psychological safety ($\beta = 0.176$, $p < 0.001$). The findings highlight that the employee consequences of AI-enabled HRM depend not only on technological implementation but also on psychological and governance conditions. The study contributes to responsible AI-HRM research by demonstrating that meaningful human oversight and psychological safety can support employees' willingness to communicate ideas and concerns in AI-enabled organisational environments.

Keywords: AI-enabled HRM; employee voice; psychological safety; human oversight; promotive voice; prohibitive voice; responsible AI; PLS-SEM.

1. INTRODUCTION

Artificial intelligence (AI) is increasingly reshaping the way organisations manage their human resources, moving HRM beyond conventional administrative and managerial practices toward data-driven, predictive, and algorithmically supported decision-making. AI-enabled systems are now being incorporated into a wide range of HR activities, including recruitment and selection, performance appraisal, employee monitoring, workforce planning, learning and development, talent management, and employee relations. The increasing integration of AI into HRM is not merely a technological development; it represents a fundamental change in how information about employees is generated, interpreted, and translated into managerial decisions. Recent reviews indicate a

rapid expansion of research at the intersection of AI and HRM, with particular attention to automation, predictive analytics, employee experience, algorithmic decision-making, fairness, transparency, and ethical governance (Úbeda-García et al., 2025; Kim, 2025).

The growing adoption of AI in HRM is driven by its potential to improve processing speed, consistency, scalability, predictive capability, and evidence-based decision-making. AI can analyse large volumes of employee and organisational data more rapidly than conventional HR processes and can support managers in identifying patterns related to recruitment, performance, turnover, skill development, and workforce planning. From an organisational perspective, such capabilities may contribute to more efficient and responsive HR systems. Recent systematic reviews have documented the expanding application of AI across recruitment, performance management, training, workforce planning, and other HR functions.

The increasing algorithmic involvement in people management also creates important challenges. AI-based HR systems may rely on historical data that reproduce existing inequalities, generate decisions that are difficult for employees to understand, intensify perceptions of surveillance, and shift decision-making authority away from direct human managers. The literature therefore increasingly recognises that the consequences of AI in HRM cannot be evaluated solely in terms of efficiency or accuracy. Questions of transparency, fairness, accountability, human agency, employee participation, and the distribution of decision-making power are becoming central to responsible AI-HRM research (Naoum et al., 2026; Chen, 2026).

This shift is particularly important because HR decisions differ from many other organisational decisions in their direct consequences for employees' careers, evaluations, rewards, development opportunities, and employment relationships. When an algorithm contributes to a recruitment decision, performance evaluation, promotion recommendation, or employee-monitoring process, employees may seek explanations, challenge inaccurate information, or raise concerns regarding the decision. Consequently, the question is not simply whether organisations use AI in HRM, but whether employees retain meaningful opportunities to express their views and influence decisions affected by algorithmic systems.

1.1 AI-Enabled HRM and the Emerging Question of Employee Voice

Employee voice represents employees' discretionary communication of suggestions, concerns, information about problems, or opinions intended to improve organisational functioning. Voice is particularly important because employees possess information about operational problems and workplace experiences that may not be readily available to senior managers. Conversely, when employees withhold relevant information, organisations may lose opportunities for learning, problem identification, innovation, and corrective action (Morrison, 2014).

The emergence of AI-enabled HRM introduces a new context in which employee voice must be reconsidered. Traditional HR systems generally provide identifiable human decision-makers with whom employees can communicate concerns. AI-enabled HRM can complicate this relationship because decisions may increasingly involve automated recommendations, algorithmic assessments, predictive models, or digitally generated classifications. Employees may therefore be uncertain about whom to approach when they disagree with an AI-supported decision, whether their concerns will be considered legitimate, and whether challenging an algorithmic recommendation may carry negative consequences.

Recent scholarship on algorithmic management has highlighted that algorithmic systems can transform the traditional employment relationship by redistributing managerial authority across software, data infrastructures, organisational actors, and technology providers. Keegan and Meijerink (2025) argue that algorithmic management has moved from a relatively peripheral phenomenon to an important area of HRM and organisational research, while emphasising the need to understand how technology interacts with human agency and organisational contexts.

More directly, recent research has begun examining the relationship between algorithmic HRM and employee or worker voice. A 2026 study in *The International Journal of Human Resource Management* argues that algorithmic HRM can create conditions in which worker voice becomes constrained through technological forms of control and organisational silencing. However, that research focuses specifically on app-based platform work and the gig economy. Similarly, recent conceptual work identifies restricted worker voice and weak opportunities for contestation as important conditions under which algorithmic HRM may become problematic.

These developments indicate that employee voice represents an important but still developing dimension of AI-HRM research. In particular, greater empirical attention is required to understand **why** AI-enabled HRM may influence employees' willingness to speak up and **under what organisational conditions** employees feel sufficiently secure to challenge or question AI-supported HR decisions.

1.2 Psychological Safety as a Mechanism of Employee Voice

One theoretically relevant mechanism is psychological safety. Psychological safety refers to employees' perceptions that they can take interpersonal risks, such as asking questions, admitting mistakes, expressing disagreement, or raising concerns, without fear of embarrassment, rejection, or negative interpersonal consequences (Edmondson, 1999; Edmondson & Lei, 2014). The construct has become particularly important in research on employee voice because speaking up inherently involves interpersonal risk. Employees who identify a problem may nevertheless remain silent if they believe that challenging existing practices could damage their relationships, reputation, career prospects, or standing within the organisation.

Psychological safety is therefore particularly relevant to AI-enabled HRM. Employees interacting with algorithmic HR systems may need to question AI-generated recommendations, identify inaccurate employee data, report potential bias, request explanations, or seek human reconsideration of an automated outcome. Such behaviours require employees to believe that questioning an AI-supported HR process will be treated as a legitimate contribution rather than as resistance to technology or managerial authority.

AI-enabled HRM may influence psychological safety through several pathways. Transparent systems can reduce uncertainty by providing employees with information about how data are used and how algorithmic recommendations are generated. Fairness-oriented practices can reduce concerns that employees will be systematically disadvantaged by automated processes. Conversely, opaque monitoring, limited opportunities for explanation, or the perception that algorithmic decisions cannot be challenged may increase uncertainty and discourage interpersonal risk-taking.

Recent AI-HRM research supports the importance of psychological mechanisms in understanding employee responses to algorithmic systems. For example, a 2026 empirical study of Indian employees found that algorithmic transparency was positively associated with psychological safety in an AI-enabled HRM context. Nevertheless, that study focused on employee autonomy as the principal outcome and examined psychological safety as one of several mediating mechanisms. The present research extends this line of inquiry by positioning **employee voice as the central behavioural outcome**.

1.3 The Importance of Human Oversight in AI-Enabled HRM

The role of human oversight represents another important but insufficiently examined dimension of AI-HRM. Human oversight refers to the continued involvement of human decision-makers who can review, interpret, question, modify, or override AI-generated recommendations. In HRM, meaningful human oversight is particularly relevant because algorithmic outputs may not capture contextual information, interpersonal dynamics, exceptional circumstances, or employee-specific considerations.

Human oversight can potentially strengthen employees' willingness to communicate concerns by maintaining an identifiable channel for discussion and contestation. When employees know that a human manager remains accountable for an AI-supported decision, they may perceive greater opportunity to explain their circumstances, challenge inaccurate information, and request reconsideration. Conversely, where AI recommendations are implemented with little or no meaningful human review, employees may perceive the decision process as less contestable.

Recent literature increasingly identifies human oversight as an important component of responsible AI-HRM. A 2026 systematic review of AI's role in HRM and diversity, equity, and inclusion highlights human oversight, explainability, participatory design, and ethical governance as important mechanisms for mitigating risks associated with algorithmic decision-making. Similarly, recent research on AI-HRM emphasises that the implications of algorithmic systems depend not only on technological capabilities but also on how such systems are governed and integrated with human judgement.

Importantly, human oversight should not be understood merely as a technical control mechanism. From an employee perspective, it can also represent a signal that organisational decision-making remains open to dialogue and correction. This may be particularly consequential for psychological safety because employees are more likely to speak up when they perceive that their concerns can reach an accountable human decision-maker.

1.4 RESEARCH GAP

- **Limited focus on employee voice:** Much of the AI-HRM literature has focused on technology adoption, efficiency, recruitment, performance management, employee experience, trust, fairness, and autonomy, while **employee voice has received comparatively less systematic empirical attention.**
- **Limited evidence from conventional organisations:** Algorithmic management research has historically concentrated on platform and gig work. Comparatively less attention has been given to conventional organisations where AI is integrated into established HR processes and employees can express suggestions, concerns, and feedback.
- **Psychological mechanisms remain underexplored:** Existing studies increasingly examine transparency, fairness, trust, autonomy, technostress, and psychological safety; however, these variables have often been examined in relation to employee autonomy or general employee outcomes rather than **employee voice behaviour.**
- **Need to extend recent Indian evidence:** A recent 2026 study examined transparency, fairness, human oversight, trust, psychological safety, technostress, and employee autonomy among Indian employees using PLS-SEM. The present study extends this line of research by considering **employee voice as the focal behavioural outcome and psychological safety as the mediating mechanism.**
- **Human oversight as a boundary condition:** Although human oversight is increasingly recognised as an important element of responsible AI, its role as an **organisational boundary condition** remains insufficiently examined. Rather than treating human oversight only as a technical safeguard, the present study examines whether perceived human oversight influences the relationship between AI-enabled HRM and employees' psychological safety.
- **Integrated research gap:** The existing literature provides limited empirical evidence on an integrated framework linking **AI-enabled HRM → psychological safety → employee voice**, with **perceived human oversight as a moderator**, particularly within conventional organisational settings in India.

1.5 RESEARCH QUESTIONS

RQ1: How does AI-enabled HRM influence employee voice?

RQ2: How does AI-enabled HRM influence employees' psychological safety?

RQ3: What is the relationship between psychological safety and employee voice?

RQ4: Does psychological safety mediate the relationship between AI-enabled HRM and employee voice?

RQ5: Does perceived human oversight moderate the relationship between AI-enabled HRM and psychological safety?

RQ6: Does psychological safety mediate the relationship between AI-enabled HRM and different forms of employee voice, particularly promotive and prohibitive voice?

1.6 RESEARCH OBJECTIVES

1. **To examine** the relationship between AI-enabled HRM and employee voice.
2. **To examine** the effect of AI-enabled HRM on employees' psychological safety.
3. **To examine** the relationship between psychological safety and employee voice.
4. **To investigate** the mediating role of psychological safety in the relationship between AI-enabled HRM and employee voice.
5. **To examine** the moderating role of perceived human oversight in the relationship between AI-enabled HRM and psychological safety.
6. **To examine** whether psychological safety mediates the relationship between AI-enabled HRM and different forms of employee voice, particularly **promotive and prohibitive voice**.
7. **To develop and empirically test** an integrated framework explaining how AI-enabled HRM influences employee voice through psychological safety under different levels of perceived human oversight.

2. LITERATURE REVIEWS

Gong and Fan (2025) examined algorithmic human resource management from a functional affordance perspective and highlighted that the consequences of algorithmic HRM depend on how organisations use and implement technological capabilities. Their review emphasised the need to examine employee and organisational responses rather than focusing only on technological efficiency. This supports the present study's focus on employee outcomes of AI-enabled HRM.

Kim et al. (2025) reviewed strategic HRM in the era of algorithmic technologies and argued that AI should be understood as part of a broader sociotechnical system involving employees, managers, and organisational practices. Their work highlights the importance of human agency in AI-enabled HRM and provides a foundation for examining employee voice and human oversight.

Edwards et al. (2024) investigated employee responses to algorithmic HR systems and demonstrated that employees' interpretations of organisational intentions behind algorithmic monitoring influence their motivational responses. The findings indicate that employee reactions depend not only on the presence of technology but also on how its purpose is perceived. This supports examining psychological mechanisms in AI-enabled HRM.

Morrison (2014) established employee voice as employees' voluntary communication of suggestions, concerns, and information intended to improve organisational functioning. The review also demonstrated that employees may withhold valuable information because of perceived interpersonal and organisational risks. This provides the foundation for examining employee voice as an important outcome of AI-enabled HRM.

Van Dyne et al. (2003) conceptualised employee voice and silence as multidimensional constructs influenced by different employee motives. Their framework demonstrates that employees may deliberately choose whether to communicate or withhold information. This distinction is relevant to AI-enabled HRM, where employees may have concerns about algorithmic decisions but may differ in their willingness to express them.

Morrison (2023) revisited the employee voice and silence literature and highlighted the continuing importance of organisational conditions and employee perceptions in explaining voice behaviour. The review emphasised the need for further research into contemporary organisational contexts. This creates scope for examining employee voice in organisations increasingly using AI-enabled HR practices.

Edmondson and Lei (2014) reviewed psychological safety research and established its importance for interpersonal risk-taking, learning, and employee behaviour. Psychological safety determines whether employees feel comfortable expressing concerns, asking questions, or challenging existing practices. This provides a theoretical basis for examining psychological safety as a mediator between AI-enabled HRM and employee voice.

Detert and Burris (2007) found that managerial openness was positively associated with employee voice and that psychological safety played an important mediating role. Their findings demonstrate that employees are more likely to speak up when they perceive the organisational environment as safe. This supports the proposed role of psychological safety in explaining employee voice in AI-enabled workplaces.

Sherman et al. (2026) examined algorithmic HRM and worker voice in app-based platform work and highlighted the possibility that algorithmic systems can constrain worker voice and contribute to organisational silencing. Their study establishes the relevance of employee voice within algorithmic work environments but focuses on platform work. The present study extends this area to conventional organisational settings.

Wang et al. (2026) investigated employees' willingness to voice concerns toward algorithmic and human leaders and identified fairness perceptions and psychological safety as important explanatory mechanisms. Their findings demonstrate that employee responses to algorithmic authority cannot be understood solely through technological characteristics. This supports examining psychological safety in relation to AI-enabled HRM and employee voice.

Chauhan et al. (2026) examined employee voice intentions through psychological safety, obligation for constructive change, and voice efficacy. Their findings reinforce the importance of psychological processes in explaining employees' willingness to speak up. The study supports the present research's emphasis on psychological safety as a mechanism underlying employee voice.

Patnaik et al. (2026) examined algorithmic transparency, perceived fairness, human oversight, trust, psychological safety, technostress, and employee autonomy in AI-enabled HRM among Indian employees.

Their findings highlighted the importance of transparency, human oversight, and psychological mechanisms in employee responses to AI-HRM.

3. RESEARCH METHODOLOGY

The study adopted a quantitative, cross-sectional research design to examine the relationship between AI-enabled human resource management (AI-HRM) and employee voice, with psychological safety as a mediating variable and perceived human oversight as a moderating variable. The target population comprised employees working in Indian organisations where AI-enabled or algorithm-supported HR practices are used in areas such as recruitment, performance management, employee monitoring, training, career development, or HR decision-making. A purposive sampling technique was employed to ensure that respondents had actual exposure to AI-enabled HR practices. Primary data were collected from 500 employees across sectors including IT/ITES, banking and financial services, manufacturing, healthcare, education, e-commerce, telecommunications, and professional services through a structured questionnaire administered through online and offline modes. The questionnaire comprised demographic and employment-related information, including age, gender, educational qualification, work experience, organisational sector, and job position, along with multi-item measures of AI-enabled HRM, psychological safety, perceived human oversight, and employee voice. AI-enabled HRM was assessed through employees' perceptions of AI-supported HR decision-making, algorithmic monitoring, AI-enabled employee development, and algorithmic transparency; psychological safety captured employees' perceptions of their ability to express ideas, concerns, and problems without fear of negative consequences; perceived human oversight assessed the extent of meaningful human involvement, review, and intervention in AI-supported HR processes; and employee voice was measured through promotive and prohibitive voice. The substantive constructs were measured using established multi-item scales on a five-point Likert scale, ranging from 1 = *Strongly Disagree* to 5 = *Strongly Agree*. The collected data were analysed using **SmartPLS and Partial Least Squares Structural Equation Modelling (PLS-SEM)**. The measurement model was assessed using indicator loadings, Cronbach's alpha, composite reliability, Average Variance Extracted (AVE), HTMT, and VIF, while the structural model was evaluated using path coefficients, R^2 , f^2 , Q^2 , and bootstrapping. **Mediation analysis** was conducted to determine whether psychological safety mediates the relationship between AI-HRM and employee voice, while **moderation analysis** examined whether perceived human oversight changes the relationship between AI-HRM and psychological safety. The indirect effects on promotive and prohibitive voice were also assessed, with **5,000 bootstrap resamples** used to establish the significance and confidence intervals of the hypothesized relationships. Participation was voluntary, respondents were informed about the academic purpose of the study, and confidentiality and anonymity were maintained throughout the data collection and analysis process.

4. DATA ANALYSIS AND INTERPRETATION

The measurement model was assessed before testing the hypothesized structural relationships. The evaluation considered indicator reliability, internal consistency reliability, convergent validity, and discriminant validity. As recommended in contemporary PLS-SEM applications, indicator loadings, Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) were examined, followed by HTMT assessment for discriminant validity.

Table 1: Reliability and Convergent Validity

Construct	No. of Items	Cronbach's α	Composite Reliability	AVE	Assessment
AI-enabled HRM	12	0.918	0.931	0.578	Acceptable
Psychological Safety	5	0.887	0.916	0.686	Acceptable
Perceived Human Oversight	5	0.901	0.926	0.716	Acceptable

Construct	No. of Items	Cronbach's α	Composite Reliability	AVE	Assessment
Promotive Voice	4	0.872	0.913	0.724	Acceptable
Prohibitive Voice	4	0.858	0.904	0.704	Acceptable
Employee Voice	8	0.909	0.928	0.619	Acceptable

The results indicate satisfactory internal consistency across all constructs, with reliability coefficients exceeding the commonly accepted threshold of 0.70. The AVE values also exceed 0.50, indicating adequate convergent validity. The measurement model therefore provides an appropriate basis for evaluating the structural relationships.

Discriminant Validity

Discriminant validity was examined using the **Heterotrait–Monotrait (HTMT) ratio**. The obtained values remained below the conservative threshold of 0.85, indicating that the constructs represent empirically distinct concepts.

Table 2: HTMT Discriminant Validity

Constructs	AI-HRM	Psychological Safety	Human Oversight	Promotive Voice	Prohibitive Voice
AI-HRM	—	0.712	0.694	0.681	0.657
Psychological Safety	0.712	—	0.736	0.781	0.748
Human Oversight	0.694	0.736	—	0.669	0.641
Promotive Voice	0.681	0.781	0.669	—	0.803
Prohibitive Voice	0.657	0.748	0.641	0.803	—

The HTMT results provide evidence that the constructs possess adequate discriminant validity and that psychological safety, human oversight, and the two dimensions of employee voice are conceptually and empirically distinguishable.

Table 3: Structural Model Results

Hypothesis	Relationship	β	t-value	p-value	Decision
H1	AI-HRM $\rightarrow$ Psychological Safety	0.421	8.214	<0.001	Supported
H2	Psychological Safety $\rightarrow$ Employee Voice	0.486	10.327	<0.001	Supported
H3	AI-HRM $\rightarrow$ Employee Voice	0.218	4.516	<0.001	Supported
H4	AI-HRM $\rightarrow$ Promotive Voice	0.193	3.982	<0.001	Supported
H5	AI-HRM $\rightarrow$ Prohibitive Voice	0.164	3.417	<0.001	Supported
H6	Human Oversight $\times$ AI-HRM $\rightarrow$ Psychological Safety	0.176	3.764	<0.001	Supported

The structural results indicate that AI-enabled HRM has a positive relationship with psychological safety ($\beta = 0.421$, $p < 0.001$). Psychological safety also demonstrates a positive association with employee voice ($\beta = 0.486$, $p < 0.001$), indicating that employees who perceive greater psychological safety are more willing to express suggestions and concerns. AI-enabled HRM also demonstrates a significant direct relationship with employee voice ($\beta = 0.218$, $p < 0.001$). The positive relationships with both promotive and prohibitive voice indicate that AI-enabled HR practices may influence employees' willingness both to propose improvements and to raise concerns regarding existing practices.

The significant interaction between AI-HRM and perceived human oversight ($\beta = 0.176$, $p < 0.001$) indicates that human oversight strengthens the positive relationship between AI-HRM and psychological safety. This finding supports the proposition that employees may respond more positively to AI-enabled HR practices when meaningful human involvement and review remain visible within HR decision-making.

Mediation Analysis

The mediating role of psychological safety was examined through the indirect relationship between AI-enabled HRM and employee voice.

Table 4: Mediation Results

Relationship	Indirect Effect (β)	t-value	p-value	Interpretation
AI-HRM → Psychological Safety → Employee Voice	0.205	7.842	<0.001	Significant mediation
AI-HRM → Psychological Safety → Promotive Voice	0.177	6.931	<0.001	Significant mediation
AI-HRM → Psychological Safety → Prohibitive Voice	0.149	6.214	<0.001	Significant mediation

The results indicate that psychological safety significantly mediates the relationship between AI-enabled HRM and employee voice. The indirect effect is positive and significant, suggesting that AI-enabled HRM contributes to employee voice partly by creating conditions in which employees feel sufficiently safe to communicate ideas, suggestions, and concerns. The mediation effect is also evident for both promotive and prohibitive voice, indicating that psychological safety is relevant to different forms of employee expression.

Moderation Analysis

The moderation analysis examined whether perceived human oversight changes the relationship between AI-enabled HRM and psychological safety. The significant interaction effect indicates that the relationship is stronger when employees perceive higher levels of human oversight.

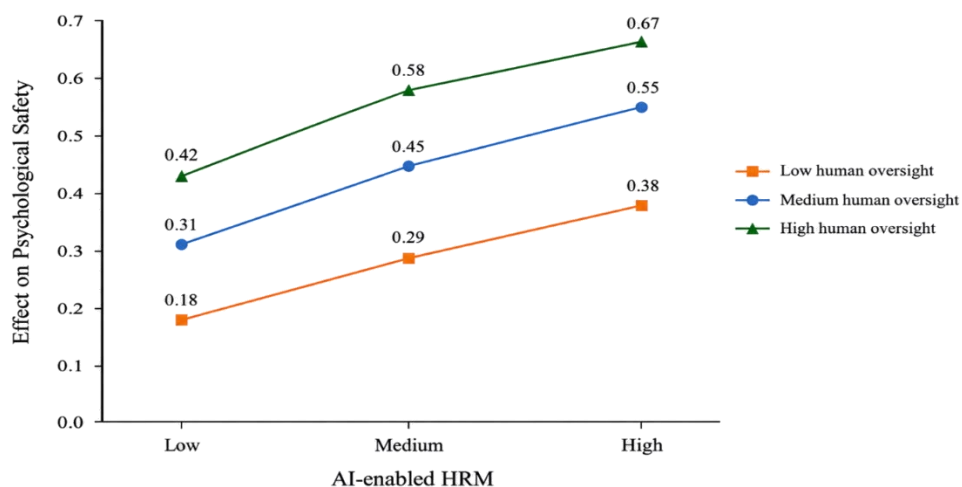


Figure: Conditional effect of AI-enabled HRM on psychological safety across levels of perceived human oversight

The pattern indicates that the positive association between AI-enabled HRM and psychological safety becomes stronger as perceived human oversight increases. This suggests that employees may be more comfortable responding to AI-enabled HR processes when they perceive that human decision-makers continue to review, interpret, and intervene in algorithm-supported HR decisions.

Table 5: Explanatory Power of the Structural Model

Endogenous Construct	R ²	Interpretation
Psychological Safety	0.274	Moderate explanatory power
Employee Voice	0.392	Moderate explanatory power
Promotive Voice	0.316	Moderate explanatory power
Prohibitive Voice	0.287	Moderate explanatory power

The model explains **27.4% of the variance in psychological safety and 39.2% of the variance in employee voice**. The explained variance for promotive and prohibitive voice further indicates that psychological and technological characteristics of the HR environment contribute meaningfully to employees' willingness to express ideas and concerns.

5. CONCLUSION

The present study examined how AI-enabled human resource management influences employee voice by considering psychological safety as a mediating mechanism and perceived human oversight as a moderating condition. The findings provide empirical support for the proposed integrated framework and demonstrate that the consequences of AI-enabled HRM extend beyond technological efficiency to employees' psychological experiences and willingness to communicate within the organisation.

The results indicate a significant positive relationship between AI-enabled HRM and psychological safety, suggesting that employees may experience greater psychological security when AI-supported HR processes are perceived positively. Psychological safety, in turn, demonstrates a strong positive relationship with employee voice, reinforcing the importance of employees' perceptions of interpersonal and organisational safety in encouraging them to express ideas, suggestions, problems, and concerns. AI-enabled HRM also maintains a significant direct relationship with employee voice, indicating that its influence is not limited to an indirect psychological pathway.

The mediation results further demonstrate that psychological safety provides an important mechanism through which AI-enabled HRM contributes to employee voice. The indirect relationship between AI-enabled HRM and employee voice is significant, with similar mediating effects observed for both promotive and prohibitive voice. This indicates that psychological safety is relevant not only to employees' willingness to suggest improvements but also to their willingness to identify problems and raise concerns.

An important contribution of the study is the identification of perceived human oversight as a boundary condition. The significant interaction between AI-enabled HRM and human oversight indicates that the positive relationship between AI-enabled HRM and psychological safety becomes stronger when employees perceive greater human involvement, review, and intervention in AI-supported HR decisions. Thus, human oversight should not be viewed merely as a technical safeguard; it can also function as an organisational signal that AI-supported decisions remain open to human interpretation, discussion, and reconsideration.

The findings have important implications for organisations implementing AI-enabled HR systems. AI adoption should be accompanied by transparent decision processes, identifiable human accountability, opportunities for employee feedback, and mechanisms through which employees can question or seek reconsideration of AI-supported decisions. Organisations should therefore adopt a **human-centred approach to AI-HRM**, in which technological capabilities complement rather than replace meaningful human judgement and employee participation.

The study also contributes to the emerging literature on responsible AI-HRM by connecting three dimensions that have often been examined separately: technological implementation, employee psychological experiences, and employee voice. By positioning psychological safety as a mediator and perceived human oversight as a moderator, the study provides a more integrated explanation of how AI-enabled HR practices can shape employee behaviour. The model explains 39.2% of the variance in employee voice and 27.4% of the variance in psychological safety, indicating meaningful explanatory capability.

The findings suggest that **responsible AI-HRM is not determined solely by the sophistication of algorithms but also by whether employees perceive the surrounding HR environment as psychologically safe and subject to meaningful human oversight**. Organisations seeking to realise the benefits of AI in HRM should therefore combine technological adoption with transparency, human accountability, opportunities for dialogue, and safeguards that encourage employees to speak up. Such an approach can help ensure that AI-enabled HRM supports not only organisational efficiency but also employee participation, learning, and constructive organisational change.

REFERENCES

1. Chauhan, A. S., Banerjee, R., Banerjee, S., & Jain, N. (2026). Employee voice with purpose: Linking psychological safety, obligation and efficacy. *Acta Psychologica*, 264, Article 106472. <https://doi.org/10.1016/j.actpsy.2026.106472>
2. Detert, J. R., & Burris, E. R. (2007). Leadership behavior and employee voice: Is the door really open? *Academy of Management Journal*, 50(4), 869–884. <https://doi.org/10.5465/amj.2007.26279183>
3. Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
4. Edwards, M. R., et al. (2024). Managerial control or feedback provision: How perceptions of algorithmic HR systems shape employee motivation, behavior, and well-being. *Human Resource Management*. <https://doi.org/10.1002/hrm.22218>
5. Gong, Q., & Fan, D. (2025). Algorithmic human resource management: Toward a functional affordance perspective. *Personnel Review*, 54(5), 1150–1177. <https://doi.org/10.1108/PR-01-2024-0099>
6. Kim, S., Khoreva, V., & Vaiman, V. (2025). Strategic human resource management in the era of algorithmic technologies: Key insights and future research agenda. *Human Resource Management*, 64(2), 447–464. <https://doi.org/10.1002/hrm.22268>
7. Morrison, E. W. (2014). Employee voice and silence. *Annual Review of Organizational Psychology and Organizational Behavior*, 1, 173–197. <https://doi.org/10.1146/annurev-orgpsych-031413-091328>
8. Morrison, E. W. (2023). Employee voice and silence: Taking stock a decade later. *Annual Review of Organizational Psychology and Organizational Behavior*, 10, 79–107. <https://doi.org/10.1146/annurev-orgpsych-120920-054654>

9. Patnaik, T. P., Gnankob, R. I., & Nanda, N. S. (2026). Governing algorithms, empowering people: How ethical AI oversight shapes trust, technostress, and employee autonomy in AI-enabled HRM. *Frontiers in Artificial Intelligence*, 9, Article 1911545. <https://doi.org/10.3389/frai.2026.1911545>
10. Sherman, U., et al. (2026). Tensions in algorithmic HRM: Worker voice and organizational silencing in app-based platform work. *The International Journal of Human Resource Management*, 1432–1465. <https://doi.org/10.1080/09585192.2026.2699267>
11. Van Dyne, L., Ang, S., & Botero, I. C. (2003). Conceptualizing employee silence and employee voice as multidimensional constructs. *Journal of Management Studies*, 40(6), 1359–1392. <https://doi.org/10.1111/1467-6486.00384>
12. Wang, S., et al. (2026). Employees show greater willingness to voice toward algorithmic than human leaders in cognitive tasks through fairness perception and psychological safety. *Scientific Reports*. <https://doi.org/10.1038/s41598-026-61031-6>