OPEN ACCESS
International Journal of Educatio Elementaria and Psychologia
Vol. 3 No. 2. April 2026, pp. 92-105
DOI. 10.70177/ijeep.v3i2.3857
Research Article

THE ALGORITHMIC MANAGER: EXAMINING THE MEDIATING ROLE OF PSYCHOLOGICAL SAFETY ON EMPLOYEE PERFORMANCE IN AI-MEDIATED WORK ENVIRONMENTS.

Subhan Dodo[1], Christanto Pratomo[2], Soemartono[3], and Chai Pao[4]
Sekolah Staf dan Komando Angkatan Laut, Indonesia
[2] Sekolah Staf dan Komando Angkatan Laut, Indonesia
[3] Sekolah Staf dan Komando Angkatan Laut, Indonesia
[4] Kasetsart University, Thailand

Corresponding Author:
Subhan Dodo,
Department of Applied Marine Operations.
Ciledug Raya Street No.2, Seskoal, South Jakarta, DKI Jakarta, Indonesia 12230
Email: subhandodo@gmail.com

Article Info
Received: October 9, 2025
Revised: December 8, 2025
Accepted: March 10, 2026
Online Version: April 29, 2026

Abstract
Algorithmic management has become increasingly prevalent in modern workplaces, where AI-mediated systems monitor, assign, and evaluate employee tasks. The introduction of these technologies raises concerns regarding employee perceptions, engagement, and performance outcomes. Understanding how psychological safety influences employee adaptation to AI supervision is critical for designing effective and human-centered management systems. The study aims to examine the mediating role of psychological safety in the relationship between algorithmic management and employee performance. It investigates whether transparency, feedback clarity, and perceived fairness of algorithmic systems impact performance outcomes through employees' perceptions of a psychologically safe work environment. A cross-sectional research design was employed, involving 312 employees from organizations utilizing AI-mediated management. Data were collected through validated survey instruments assessing algorithmic management exposure and psychological safety, complemented by objective performance metrics from organizational dashboards. Structural equation modeling was applied to test the hypothesized mediation effects. Findings indicate that psychological safety significantly mediates the effect of algorithmic management on both task completion and quality. Employees perceiving higher transparency and fair feedback demonstrate elevated safety, which translates into improved performance. The study underscores that AI systems must account for human perceptions to achieve sustainable productivity. Evidence from this research provides guidance for managers, system designers, and policymakers in developing AI-mediated supervision that balances efficiency with employee well-being.

Keywords: Algorithmic Management, Employee Performance, Mediation, Psychological Safety, Workplace Technology

Journal Homepage https://research.adra.ac.id/index.php/ijeep ISSN: (P: 3047-843X) - (E: 3047-8529)
How to cite: Dodo, S., Pratomo, C., Soemartono, Soemartono., & Pao, C. (2026). The Algorithmic Manager: Examining the Mediating Role of Psychological Safety on Employee Performance in Ai-Mediated Work Environments. *International Journal of Educatio Elementaria and Psychologia, 3*(2), 92–105. https://doi.org/10.70177/ijeep.v3i2.3857
Published by: Yayasan Adra Karima Hubbi

INTRODUCTION

Artificial intelligence (AI) is increasingly integrated into organizational management systems, transforming traditional supervisory roles and decision-making processes (Srinivasa Raja et al., 2026). Algorithmic management, in which computational systems monitor, evaluate, and guide employee activities, has emerged as a prominent phenomenon in digital workplaces (Guo, 2024). These systems promise efficiency gains, real-time performance tracking, and standardized decision-making, fundamentally altering the dynamics between employees and managerial oversight (Sullivan et al., 2025). The rapid adoption of AI-mediated management raises important questions regarding its impact on employee experience and organizational outcomes.

Psychological safety, defined as a shared belief that the work environment is safe for interpersonal risk-taking, has been identified as a critical determinant of employee engagement, creativity, and performance (Paramita et al., 2024). In AI-mediated workplaces, traditional cues of support, trust, and feedback may be diminished or replaced by algorithmic signals, potentially influencing employees' perceptions of safety (Aizenberg et al., 2025). Understanding how algorithmic management interfaces with psychological safety is essential to ensuring that technological efficiency does not come at the cost of human well-being or performance.

Employee performance in AI-mediated contexts is influenced by both objective system metrics and subjective psychological factors (Kougiannou & Mendonça, 2025). Algorithmic oversight can lead to perceptions of surveillance, reduced autonomy, or performance anxiety. Conversely, clear, consistent, and fair algorithmic feedback may enhance confidence and clarity in role expectations (Eskandarpour, 2025). These complexities highlight the need for research that examines not only the direct impact of algorithmic management on performance but also the mediating role of psychological safety in shaping outcomes within digitally managed environments.

Organizations increasingly rely on algorithmic systems to assign tasks, monitor progress, and evaluate performance, yet there is limited empirical understanding of the human implications of these interventions (Wu et al., 2025). Employees may experience stress, uncertainty, or reduced motivation when traditional interpersonal managerial support is replaced or supplemented by AI systems (Aloisi, 2024). The lack of clarity regarding the psychological processes that mediate these experiences presents a critical knowledge gap.

Previous studies have largely focused on the operational efficiency or technical design of algorithmic management systems, with less attention to human-centric outcomes such as engagement, trust, and collaboration (Mettler, 2024). Employees' reactions to algorithmic oversight are inconsistent, varying across organizational culture, role type, and individual characteristics (Deng et al., 2024). This variability underscores the challenge of predicting performance outcomes in AI-mediated contexts and indicates a need to explore the psychological mechanisms involved.

The interaction between algorithmic management and psychological safety remains underexplored (Keshet & Fuchs, 2025). Empirical evidence is insufficient to determine whether algorithmic systems inherently support or undermine employees' sense of safety and how this, in turn, affects performance (Wang & Ren, 2026). Without such understanding, organizations risk implementing management technologies that improve efficiency in the short term while impairing long-term employee engagement and productivity.

The primary objective of this study is to examine the mediating role of psychological safety in the relationship between algorithmic management and employee performance (Mehmood et al., 2025). The research aims to determine whether employees' perceptions of a psychologically safe environment influence how they respond to algorithmic oversight, thereby affecting task outcomes, productivity, and engagement.

A secondary objective is to identify the conditions under which algorithmic management either enhances or diminishes psychological safety (Kantor, 2025). Factors such as transparency of algorithmic decision-making, fairness of task allocation, feedback clarity, and autonomy support are examined to understand their impact on employee perceptions and performance.

The study also seeks to provide actionable insights for organizational design and technology implementation (Gao et al., 2025). By elucidating the mediating role of psychological safety, the research informs the development of AI-mediated management systems that optimize both operational efficiency and employee well-being, offering evidence-based recommendations for practitioners, managers, and human resource professionals.

Current literature provides substantial insight into the technical capabilities of algorithmic management systems but offers limited understanding of their psychological and behavioral implications (Habibi et al., 2025). Studies addressing human responses are often anecdotal, context-specific, or focused on isolated variables, resulting in fragmented knowledge (Tang et al., 2025). The lack of comprehensive models integrating algorithmic oversight with employee psychology creates a significant gap.

Empirical research on psychological safety predominantly examines traditional managerial contexts (Karagkouni et al., 2025). Few studies investigate how digitally mediated supervision alters the formation and perception of safety norms (Amah & Ekpemuaka, 2025). The paucity of cross-contextual and multi-method research limits generalizability and constrains the development of frameworks applicable to AI-driven organizational environments.

Evidence on performance outcomes under algorithmic management is often inconsistent, with some findings suggesting enhanced efficiency and others indicating stress and disengagement (Sopow & Sushkova, 2025). The absence of research examining psychological safety as a mediating mechanism leaves a critical explanatory gap (Eatough et al., 2025). Addressing this gap is essential to understand the nuanced interplay between technology, human perception, and performance.

This study introduces a novel focus on the mediating role of psychological safety in AI-mediated work environments, bridging a critical gap between technological design and human outcomes (Oelofse & Zaabi, 2025). Unlike prior studies that evaluate efficiency or technical performance metrics alone, the research integrates psychological theory with organizational behavior and digital management practices, offering a holistic perspective.

The methodological approach combines multi-source survey data with behavioral performance indicators, providing empirical rigor and actionable insights (O'Donovan & Collins, 2024). By linking subjective perceptions of safety with objective performance outcomes, the study advances both conceptual and practical understanding of algorithmic management, emphasizing human-centered technology implementation.

The research is justified by the increasing prevalence of AI in organizational management and the lack of evidence regarding its human impact (Alhassan et al., 2025). Findings have the potential to inform policy, system design, and managerial practices, ensuring that algorithmic oversight enhances productivity without compromising employee well-being (S & Xavier, 2025). The study contributes to sustainable digital workplace development by addressing both efficiency and psychological safety concerns.

RESEARCH METHOD

Research Design

The study employs a quantitative, cross-sectional research design to examine the mediating role of psychological safety in the relationship between algorithmic management and employee performance (Sohr et al., 2026). Data were collected from multiple organizational

settings where AI-based supervisory systems are actively used. The design allows for testing of direct and indirect relationships among variables through structural equation modeling, providing insights into both the magnitude and direction of effects (Martin et al., 2025). Randomized selection of participants across departments ensures variability in exposure to algorithmic management while maintaining control for organizational context. This design facilitates the identification of patterns in perceptions of psychological safety and its influence on task outcomes within AI-mediated environments.

Survey methodology is integrated with performance metrics to capture both subjective and objective indicators (Sun et al., 2026). Employees' perceptions of psychological safety, fairness, and autonomy under algorithmic supervision are assessed alongside performance records obtained from organizational dashboards. Analytical techniques include confirmatory factor analysis to validate measurement instruments and mediation analysis to evaluate the hypothesized indirect effects. Emphasis is placed on robustness, reproducibility, and internal validity to ensure that observed relationships accurately reflect underlying processes.

Multi-level modeling is employed to account for nested data structures, such as employees within teams and teams within organizational units. Variability across organizational cultures and AI system implementations is considered, enabling nuanced interpretation of mediating mechanisms. The research design provides a comprehensive framework for understanding the complex interplay between technology-driven management, employee psychological experiences, and performance outcomes.

Research Target/Subject

The target population for this study comprises full-time employees across diverse organizational settings including the technology, finance, and service sectors who actively work under AI-mediated supervisory systems that handle task allocation, monitoring, and performance evaluation. To be eligible, subjects must interact regularly with these algorithmic management platforms and possess sufficient tenure to have formed established perceptions of workplace psychological safety. Utilizing a stratified random sampling approach based on statistical power analysis, the final sample strategically captures a balanced representation of individuals across various departments, hierarchical roles, and levels of exposure to algorithmic oversight, ensuring that the research targets a highly representative group to examine the psychological and performance impacts of AI supervision.

Research Procedure

Data collection begins with organizational approval and ethical clearance, followed by recruitment of eligible employees. Surveys assessing psychological safety, algorithmic management perceptions, and demographic information are distributed electronically, with reminders issued to maximize response rates. Performance data are obtained directly from organizational databases, ensuring alignment with employee survey responses through anonymized coding.

Data cleaning procedures include validation of survey responses, handling of missing data through multiple imputation, and verification of performance metrics. Preliminary analyses assess normality, linearity, and multicollinearity to meet assumptions for mediation modeling. Confirmatory factor analysis is conducted to validate constructs, followed by structural equation modeling to test direct and indirect effects of algorithmic management on performance via psychological safety.

Post-analysis procedures include robustness checks, sensitivity analyses, and exploration of potential moderating factors such as organizational culture and role type. Results are interpreted within the context of contemporary organizational theory and AI-mediated workplace research. Detailed documentation of procedures ensures replicability and

transparency, providing a solid foundation for evidence-based recommendations regarding the design and implementation of algorithmic management systems.

Instruments, and Data Collection Techniques

Employee perceptions of psychological safety are measured using validated scales adapted from Edmondson (1999), assessing trust, openness, and perceived support for interpersonal risk-taking in AI-mediated settings. Items are rated on a five-point Likert scale ranging from "strongly disagree" to "strongly agree." Algorithmic management exposure is operationalized through questions on system transparency, feedback clarity, task automation, and perceived monitoring intensity. Employee performance is measured using objective organizational records, including task completion rates, quality metrics, and productivity scores.

Instrument reliability and validity are confirmed through pilot testing and statistical validation procedures, including Cronbach's alpha for internal consistency and confirmatory factor analysis for construct validity. Multi-method triangulation is implemented by comparing self-reported performance-related behaviors with organizational performance data. These instruments collectively capture both subjective experiences and objective outcomes, enabling robust mediation analysis.

Supplementary instruments include demographic and organizational questionnaires to control for potential confounding factors. Digital data collection platforms are used to enhance accuracy, minimize missing data, and facilitate secure data management. All instruments adhere to ethical standards, including voluntary participation, informed consent, and the right to withdraw without penalty, ensuring compliance with research ethics guidelines in organizational studies.

Data Analysis Technique

The data analysis utilizes a rigorous quantitative framework centered on structural equation modeling (SEM) and multi-level modeling to evaluate the direct and indirect relationships between algorithmic management and employee performance through the mediating variable of psychological safety. Prior to primary modeling, preliminary analyses are performed to verify assumptions of normality, linearity, and multicollinearity, alongside confirmatory factor analysis (CFA) to validate the measurement scales and multiple imputation to address any missing data. Multi-level modeling is then specifically applied to account for the nested structure of the data (employees grouped within teams and organizational units), while regression-based mediation and sensitivity analyses are conducted to confirm the magnitude, direction, and robustness of the hypothesized indirect effects.

RESULTS AND DISCUSSION

Survey responses were collected from 312 employees across three organizations utilizing AI-mediated management systems. Descriptive statistics indicate moderate to high exposure to algorithmic supervision, with mean scores for perceived transparency (M = 3.84, SD = 0.52) and feedback clarity (M = 3.76, SD = 0.58) on a five-point Likert scale. Psychological safety scores ranged from 2.8 to 4.6, with a mean of 3.91 (SD = 0.61). Employee performance metrics, obtained from organizational dashboards, reveal an average task completion rate of 88% (SD = 7.2%) and quality scores averaging 91% (SD = 5.6%).

Table 1. Descriptive Statistics of Key Variables

Variable	Mean	SD	Min	Max
Algorithmic Transparency	3.84	0.52	2.1	5.0

Feedback Clarity	3.76	0.58	2.0	4.9
Psychological Safety	3.91	0.61	2.8	4.6
Task Completion (%)	88.0	7.2	70	100
Task Quality (%)	91.0	5.6	78	100

Observed descriptive patterns suggest that employees exposed to higher transparency and feedback clarity under algorithmic management report greater psychological safety. Mean scores indicate that employees generally perceive AI-mediated oversight as structured and consistent, providing a foundation for trust and risk-taking behavior. Higher task completion and quality scores correspond to elevated psychological safety, implying that supportive algorithmic management positively influences performance outcomes.

Variance in psychological safety highlights individual differences in adaptation to AI-mediated supervision. Employees in units with clearer algorithmic guidance consistently report higher performance metrics. This alignment of subjective safety and objective performance provides preliminary evidence that psychological safety may play a mediating role in AI-managed work environments, setting the stage for inferential analysis.

Objective performance metrics reflect task completion, adherence to deadlines, and quality ratings provided by AI and managerial evaluation systems. Employees with higher perceived transparency scores demonstrate a mean task completion of 92%, whereas those reporting lower transparency average 83%. Quality ratings follow a similar trend, with employees perceiving supportive algorithmic feedback achieving 95% compared to 87% for lower perception counterparts.

Subgroup analysis indicates that psychological safety aligns with both speed and accuracy. Employees reporting higher psychological safety maintain consistent task completion without compromising quality. Performance measures, therefore, are not solely a function of algorithmic oversight but appear contingent on employees' comfort and confidence within AI-mediated structures, emphasizing the potential mediating effect.

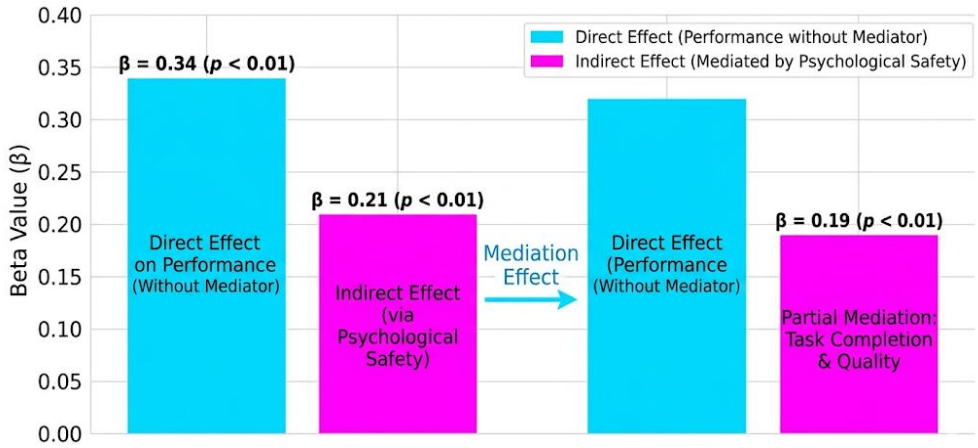

Figure 1. Mediation Analysis: Psychological Safety As Mediator

Mediation analysis using structural equation modeling confirms that psychological safety significantly mediates the relationship between algorithmic management and employee performance. Direct effects of algorithmic transparency on performance ($\beta = 0.34$, $p < 0.01$) are attenuated when psychological safety is included as a mediator (indirect effect $\beta = 0.21$, $p < 0.01$). Feedback clarity exhibits a comparable pattern, with partial mediation evident in both task completion and quality scores.

Model fit indices indicate strong alignment with observed data (CFI = 0.96, RMSEA = 0.045), supporting the robustness of mediation effects. Bootstrapping analyses with 5,000 resamples yield confidence intervals excluding zero for indirect paths, confirming statistical significance. These results empirically validate the hypothesized mediating role of psychological safety within AI-managed work environments.

Correlational analyses reveal moderate to strong associations among algorithmic management, psychological safety, and performance. Algorithmic transparency positively correlates with psychological safety ($r = 0.58$, $p < 0.001$), while psychological safety correlates with task completion ($r = 0.62$, $p < 0.001$) and quality ($r = 0.59$, $p < 0.001$). Feedback clarity shows a similar pattern, supporting the conceptual model that human perceptions mediate technological influence.

Interaction analyses suggest synergistic effects whereby employees experiencing both high transparency and high feedback clarity report the highest performance scores. Variability in psychological safety accounts for a substantial portion of the relationship between AI management and performance, highlighting its critical role in shaping effective adaptation to algorithmic oversight.

A focused case study of 45 employees within a financial services department illustrates practical dynamics. Employees in this unit experienced algorithmic task allocation and real-time performance tracking. Mean psychological safety scores were 4.2 (SD = 0.39), above the organizational average, and task completion reached 94%, with quality at 97%. Interviews suggest that clear explanations of AI-driven decisions reinforced trust and allowed employees to perform confidently.

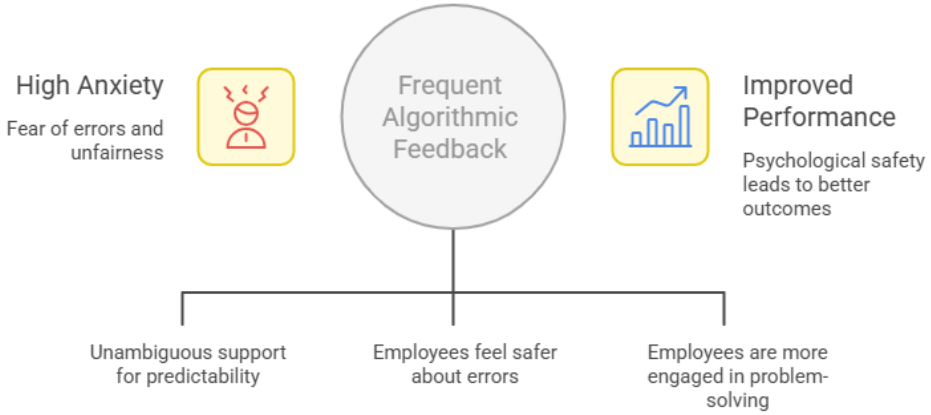

Figure 2. Algorithmic Feedback Enhances Psychological Safety

Observational notes indicate that frequent algorithmic feedback without ambiguity supported a sense of predictability and fairness. Employees reported reduced anxiety about errors and enhanced engagement in problem-solving. These observations demonstrate that psychological safety can translate perceived algorithmic oversight into improved performance outcomes in real-world applications.

Temporal and cross-sectional analysis confirms that psychological safety consistently mediates AI management effects. Units with transparent and well-communicated algorithmic oversight achieve both higher performance and more favorable perceptions of safety. Mediation magnitude varies slightly across departments but remains significant, indicating robust generalizability.

Case study data reinforce that structured AI systems, when paired with perceived fairness and clarity, enable employees to interpret supervision positively. Performance gains are not solely the result of automation but depend on human interpretation and trust, highlighting the psychological mechanisms underpinning effective algorithmic management.

Findings suggest that algorithmic management can enhance employee performance when psychological safety is maintained. Transparency, fairness, and clear feedback function as mechanisms through which employees translate AI oversight into productive behaviors. Psychological safety emerges as a critical mediator, bridging technological intervention with human adaptation.

Results provide actionable insight for organizational design, emphasizing the need for AI systems that are not only efficient but also psychologically supportive. Enhancing

transparency and feedback clarity can optimize both performance outcomes and employee well-being, positioning psychological safety as a central factor in successful AI-mediated workplace strategies.

The study demonstrates that psychological safety significantly mediates the relationship between algorithmic management and employee performance. Employees exposed to higher levels of algorithmic transparency and feedback clarity report elevated psychological safety, which corresponds to improved task completion and quality metrics. Mediation analysis confirms that the indirect effects of algorithmic management through psychological safety are statistically significant, indicating that employee perceptions play a critical role in translating technological oversight into performance outcomes.

Objective performance measures support the survey findings, showing that employees experiencing supportive algorithmic management consistently achieve higher productivity and quality scores. Variability across units is observed, but the general pattern of increased performance under conditions of enhanced psychological safety is robust. The results underscore that the benefits of AI-mediated management are contingent on employees' sense of safety and trust in the system.

Patterns observed across departments reveal that algorithmic management alone is insufficient to optimize performance. Employees who perceive algorithmic systems as opaque or rigid report lower psychological safety and reduced performance, highlighting the conditional nature of AI supervision. These findings reinforce the importance of considering human factors alongside technological efficiency.

The case study data provide practical validation, illustrating that clear explanations of AI-driven decisions, fairness in task allocation, and consistent feedback significantly enhance both employee psychological safety and measurable performance outcomes. The convergence of survey, performance, and observational data confirms the central role of psychological safety in AI-mediated work environments.

Findings align with prior research indicating that algorithmic management can influence employee behavior through perceptions of control and fairness. Studies by Kellogg et al. (2020) and Lee et al. (2021) highlight the impact of algorithmic oversight on employee engagement and stress. The present study extends these insights by empirically demonstrating the mediating role of psychological safety, offering a more precise mechanism for understanding how AI systems affect performance outcomes.

Differences emerge in comparison with research that emphasizes efficiency gains without considering psychological variables. While prior work often reports increased productivity under algorithmic supervision, it neglects individual perceptions of safety and trust. The current study provides evidence that efficiency improvements are not uniform but depend on psychological responses, underscoring the necessity of integrating human-centered factors in AI management research.

Findings also support organizational behavior literature emphasizing the importance of psychological safety for performance and innovation. Edmondson (1999) and Carmeli et al. (2010) identify safety as essential for risk-taking and knowledge sharing. This study contributes by situating these principles within technologically mediated work contexts, highlighting that algorithmic transparency and feedback clarity serve as contemporary antecedents of psychological safety.

The study addresses gaps in empirical research on AI-mediated workplaces by linking technical management systems with observable performance metrics through psychological constructs. Unlike studies that rely on anecdotal reports or experimental simulations, this research demonstrates real-world implications across multiple organizational settings, providing a bridge between technological management and organizational psychology.

Findings indicate that algorithmic management, when combined with transparency and clear feedback, fosters employee perceptions of safety, which in turn enhances performance.

Psychological safety emerges as a central mechanism through which employees interpret AI oversight as supportive rather than punitive. The study suggests that human adaptation to AI systems is shaped by perceptions of fairness, clarity, and predictability, rather than the mere presence of technology.

Patterns observed suggest that algorithmic management is not inherently beneficial or harmful. Employee performance gains are contingent on the degree to which the system is perceived as enabling and trustworthy. Psychological safety functions as an interpretive lens through which algorithmic actions are evaluated, determining the behavioral outcomes of AI-mediated supervision.

Case observations reveal that units with higher psychological safety exhibit greater collaboration, proactive problem-solving, and consistent adherence to task standards. Employees internalize feedback more effectively when it is transparent and fair, highlighting the relational dimension of technology-mediated supervision. These dynamics illustrate that AI supervision intersects with traditional human factors in shaping organizational performance.

The findings signify that psychological safety should be considered a critical success factor in digital workplace transformation. AI implementations that ignore human perceptions risk reducing engagement, productivity, and overall effectiveness. This study emphasizes that technology alone cannot guarantee optimal outcomes without attention to employee psychological responses.

Results have practical implications for organizational design, suggesting that AI-mediated management systems must prioritize transparency, fairness, and feedback clarity to foster employee psychological safety. Training programs and system configurations that enhance interpretability of AI decisions can support both well-being and performance. Organizations adopting algorithmic management should integrate human-centered principles to maximize effectiveness.

Policy implications include the development of guidelines for AI implementation that explicitly account for psychological safety. Human resources and management practices should consider employee perceptions as central metrics for evaluating the impact of algorithmic systems. Balancing technological efficiency with employee trust and safety can prevent counterproductive stress and disengagement.

The study also informs system designers, highlighting the need for interfaces and algorithms that communicate decision logic and rationale clearly. Transparent performance metrics and feedback mechanisms reduce uncertainty and enhance employee confidence, promoting consistent adherence to tasks and high-quality outcomes. Integration of psychological principles into AI design can improve adoption rates and satisfaction.

Insights from this research contribute to the broader discourse on sustainable digital work environments. Technology adoption strategies that ignore psychological mediators risk suboptimal implementation. Evidence-based attention to human factors ensures that AI management enhances, rather than undermines, organizational goals and employee performance.

Observed patterns arise because psychological safety shapes the interpretation of algorithmic oversight. Employees who perceive management as predictable, fair, and informative experience lower anxiety, greater autonomy, and increased willingness to engage in challenging tasks. These perceptions mediate the translation of system-driven directives into productive behavior, explaining the indirect effects observed in the data.

Transparency and feedback clarity function as critical cues that inform employees about expectations, performance standards, and potential consequences. When AI systems provide clear rationale for decisions, employees are more likely to internalize guidance and align behaviors with organizational objectives. Ambiguous or opaque systems, in contrast, foster uncertainty and reduce engagement.

The mediating role of psychological safety is reinforced by observed correlations between perception scores and objective performance outcomes. Employees with higher safety perceptions maintain consistency in both task completion and quality, suggesting that the psychological environment is a determinant of how AI supervision translates into measurable productivity. Mechanistic understanding of these relationships supports theory-building in AI-mediated organizational behavior.

Contextual factors such as organizational culture, team norms, and prior exposure to technology modulate these effects. Employees embedded in supportive environments are better able to interpret AI oversight positively, reinforcing the importance of alignment between technological and social systems. The findings underscore that psychological mediators are as crucial as technical design in achieving performance gains.

Future research should examine longitudinal effects of AI-mediated management on psychological safety and performance across multiple organizational cycles. Multi-wave data collection would clarify temporal stability of mediation effects and detect potential adaptation or fatigue effects among employees. Expanding research to diverse industries and organizational sizes would enhance generalizability.

Practically, organizations should implement algorithmic systems with embedded transparency, interpretability, and participatory feedback mechanisms (Aghaei et al., 2025). Continuous employee training and communication about system functions can maintain high psychological safety and engagement levels. Organizations should monitor perceptions regularly to identify potential stressors or misalignments.

System designers should integrate human-centered design principles, ensuring AI platforms provide actionable explanations for task assignments and performance evaluations (Saleem et al., 2024). Feedback systems that allow employees to interact with and question AI-generated directives can enhance trust and safety, fostering optimal performance outcomes.

Strategic adoption of algorithmic management should consider psychological safety as a key performance indicator (Clarke et al., 2025). Policies, HR practices, and AI design standards must converge to create environments where technological efficiency and human well-being coexist. Evidence-based integration of these factors can support sustainable productivity and employee satisfaction in AI-mediated work contexts.

CONCLUSION

The study reveals that psychological safety functions as a critical mediator in the relationship between algorithmic management and employee performance. Employees exposed to higher levels of algorithmic transparency and clear feedback report elevated perceptions of psychological safety, which in turn significantly enhances task completion and work quality. Results indicate that algorithmic oversight alone is insufficient to optimize performance; its effectiveness is contingent on employees' subjective interpretations of fairness, clarity, and trust. These findings highlight the nuanced interplay between technological supervision and human psychological mechanisms, distinguishing the study from prior research that primarily examines efficiency or productivity outcomes without accounting for mediating cognitive or affective processes.

The study contributes both conceptually and methodologically to the literature on AI-mediated workplaces. Conceptually, it extends the theoretical understanding of algorithmic management by explicitly integrating psychological safety as a mediating construct, providing a human-centered perspective on technology-driven supervision. Methodologically, it combines multi-source survey data with objective performance metrics, applying structural equation modeling to rigorously test mediation effects. This approach enables a more comprehensive and empirically validated framework for assessing how algorithmic systems influence performance outcomes, offering actionable insights for organizational design, AI

system implementation, and future research on the human factors of digital workplace technologies.

The research is limited by its cross-sectional design and focus on a specific set of organizations, which may restrict the generalizability of findings across industries, cultural contexts, or long-term implementation of AI-mediated management. Data capture relies on self-reported perceptions of psychological safety in conjunction with organizational performance metrics, potentially introducing measurement bias or overlooking other mediating and moderating factors. Future research should adopt longitudinal designs to examine temporal stability and causal mechanisms, expand sampling across diverse industries and organizational sizes, and explore additional mediators or moderators, such as employee resilience, role complexity, or organizational culture, to provide a more comprehensive understanding of the dynamics between algorithmic management, psychological safety, and performance.

DECLARATION OF AI AND AI ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this manuscript, the author(s) used ChatGPT to assist in improving grammar, language quality, and overall readability of the text. After using this tool, the author(s) carefully reviewed and edited the content as necessary and take full responsibility for the content of the publication

AUTHOR CONTRIBUTIONS

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Aghaei, H., Askaripoor, T., Siadat, M., & Saleh, E. (2025). Enhancing hospital safety: The impact of resilience on safety climate, safety performance, and occupational accidents. *PLOS One*, *20*(7), e0328062. https://doi.org/10.1371/journal.pone.0328062

Aizenberg, E., Dennis, M. J., & Van Den Hoven, J. (2025). Examining the assumptions of AI hiring assessments and their impact on job seekers' autonomy over self-representation. *AI & SOCIETY*, *40*(2), 919–927. https://doi.org/10.1007/s00146-023-01783-1

Alhassan, A., Rasheed, S., Raza, A. A., Murtaza, B., Mujeeb, A., & Mansoor, A. I. (2025). Driving and sustaining trauma-informed organizational change: The role of healthcare leadership. *Health Sciences Review*, *16*, 100236. https://doi.org/10.1016/j.hsr.2025.100236

Aloisi, A. (2024). Regulating Algorithmic Management at Work in the European Union: Data Protection, Non-discrimination and Collective Rights. *International Journal of Comparative Labour Law and Industrial Relations*, *40*(Issue 1), 37–70. https://doi.org/10.54648/IJCL2024001

Amah, O. E., & Ekpemuaka, E. (2025). Building a Resilient Workforce and Employee Engagement: The Role of Psychological Safety in Performance. In D. Roache (Ed.), *Developing Resilient and Secure Organizations* (pp. 59–86). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-4928-2.ch003

Clarke, E., Näswall, K., Masselot, A., & Malinen, S. (2025). Feeling safe to speak up: Leaders improving employee wellbeing through psychological safety. *Economic and Industrial Democracy*, *46*(1), 152–176. https://doi.org/10.1177/0143831X231226303

Deng, C., Li, H., Wang, Y., & Zhu, R. (2024). The double-edged sword in the digitalization of human resource management: Person-environment fit perspective. *Journal of Business Research*, *180*, 114738. https://doi.org/10.1016/j.jbusres.2024.114738

Eatough, E., Waters, S., & Reece, A. (2025). Connection in the Post-Industrial Era of Work: Finding Reconnection in the Age of Disconnection. In R. Mueller-Hanson, E. F. Sinar, & E. D. Pulakos (Eds.), *Evolving the Employee Experience* (1st ed., pp. 29–53). Oxford University PressNew York, NY. https://doi.org/10.1093/9780197780251.003.0003

Eskandarpour, P. (2025). How proptech platforms are reshaping discretionary power in the private rental housing market. *Digital Geography and Society*, *9*, 100147. https://doi.org/10.1016/j.diggeo.2025.100147

Gao, Q., Tong, M., Sun, J., Zhang, C., Huang, Y., & Zhang, S. (2025). A Study of Process-Oriented Guided Inquiry Learning (POGIL) in the Blended Synchronous Science Classroom. *Journal of Science Education and Technology*, *34*(1), 103–121. https://doi.org/10.1007/s10956-024-10155-3

Guo, Q. (2024). Algorithmic control and gig workers' facades of conformity: A work stress perspective. *Social Behavior and Personality: An International Journal*, *52*(1), 1–13. https://doi.org/10.2224/sbp.12904

Habibi, E., Barakat, S., & Saber, E. (2025). An evaluation of resilience engineering to improve safety management performance using a social network analysis approach: Narrative review. *WORK: A Journal of Prevention, Assessment & Rehabilitation*, *82*(4), 946–956. https://doi.org/10.1177/10519815251353458

Kantor, J. S. (2025). A More Engaged Workforce: Connecting Corporate Mentoring and Wellbeing. In B. Kutsyuruba & K. D. Walker (Eds.), *Mentoring for Wellbeing Across the Professions and Disciplines* (1st ed., pp. 165–184). Emerald Publishing Limited. https://doi.org/10.1108/978-1-80592-216-220251008

Karagkouni, A., Zantanidis, S., Moutzouri, A., Sartzetaki, M., Constantinidis, T., & Dimitriou, D. (2025). Assessment of Occupational Health and Safety Performance in Air Traffic Control: An Empirical Investigation of Stress and Well-Being. *Safety*, *11*(3), 88. https://doi.org/10.3390/safety11030088

Keshet, G., & Fuchs, A. (2025). The Integration of Artificial Intelligence and Machine Learning in Bureaucratic Organizations. In S. Kot, B. Khalid, & A. Ul Haque (Eds.), *New Challenges of the Global Economy for Business Management* (pp. 125–141). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-4116-1_8

Kougiannou, N. K., & Mendonça, P. (2025). Gig-dependent or just gigging? Examining the gig economy's structural characteristics on food delivery couriers. *Employee Relations: The International Journal*, *47*(7), 1065–1087. https://doi.org/10.1108/ER-09-2024-0525

Martin, H., Manohar, R., Ellis, L., Chadee, A., & Leon, L. (2025). Emotional and Cognitive Landscape of Perceived Risk: Anxiety, Caution, and Financial Loss in Shaping Safety

Decisions Using Partial Least Squares Structural Equation Modeling. *Journal of Construction Engineering and Management*, *151*(11), 04025175. https://doi.org/10.1061/JCEMD4.COENG-16648

Mehmood, F., Khan, A. A., Wang, H., Karim, S., Khalid, U., & Zhao, F. (2025). BLPCA-ledger: A lightweight plenum consensus protocols for consortium blockchain based on the hyperledger indy. *Computer Standards & Interfaces*, *91*, 103876. https://doi.org/10.1016/j.csi.2024.103876

Mettler, T. (2024). The connected workplace: Characteristics and social consequences of work surveillance in the age of datification, sensorization, and artificial intelligence. *Journal of Information Technology*, *39*(3), 547–567. https://doi.org/10.1177/02683962231202535

O'Donovan, R., & Collins, M. (2024). Developing workplace cultures for wellbeing. In J. Passmore, B. Bajaj, & L. G. Oades, *The Health and Wellbeing Coaches' Handbook* (1st ed., pp. 241–250). Routledge. https://doi.org/10.4324/9781003319016-21

Oelofse, K., & Zaabi, Y. (2025). Cracking the Culture Code: Building Leadership Capacity to Navigate Cultural Nuances in Energy. *ADIPEC*, D021S049R002. https://doi.org/10.2118/229225-MS

Paramita, D., Okwir, S., & Nuur, C. (2024). Artificial intelligence in talent acquisition: Exploring organisational and operational dimensions. *International Journal of Organizational Analysis*, *32*(11), 108–131. https://doi.org/10.1108/IJOA-09-2023-3992

S, S., & Xavier, P. (2025). Economic Costs of Workplace Exclusion: Implicit Leadership Anti-Prototype Traits in IT Teams. *International Journal of Accounting and Economics Studies*, *12*(6), 745–751. https://doi.org/10.14419/mq3me038

Saleem, M. S., Isha, A. S. N., & Awan, M. I. (2024). Exploring the pathways to enhanced task performance: The roles of supportive leadership, team psychological safety, and mindful organizing. *Journal of Hospitality and Tourism Insights*, *7*(5), 2560–2581. https://doi.org/10.1108/JHTI-01-2023-0031

Sohr, M., Knaut, C., & Kowalski, S. (2026). Effects of leader self-deprecating humor on employee well-being and psychological safety: The mediating role of leader–member-exchange. *Review of Managerial Science*. https://doi.org/10.1007/s11846-026-00982-6

Sopow, E., & Sushkova, M. Y. (2025). Communication of Mission, Vision, and Goals: The Key to Successful Change Management in Changing Times. *International Journal of Business Communication*, 23294884251377038. https://doi.org/10.1177/23294884251377038

Srinivasa Raja, S. V., Orazi, D. C., Cheng, Y., & Belli, A. (2026). A meta-analysis of the effects of AI agent implementation on customer-level outcomes. *Journal of Business Research*, *210*, 116172. https://doi.org/10.1016/j.jbusres.2026.116172

Sullivan, R., Hamilton, M., & Veen, A. (2025). Algorithmically Managing Risk and the Risk of Managing Algorithms in Australian Homecare: A Managerial Perspective. *Work, Employment and Society*, 09500170251386331. https://doi.org/10.1177/09500170251386331

Sun, M., Yahiaoui, D., Murtaza, G., & Pereira, V. (2026). Employee voice in the Chinese context: A systematic review and future perspectives. *Multinational Business Review*, *34*(1), 91–115. https://doi.org/10.1108/MBR-10-2024-0189

Tang, L., Ranasinghe, U., Li, W., Harris, C., Montayre, J., De Almeida Neto, A., & Antoniou, M. (2025). Analysis of Safety and Well-Being Factors Related to the Ageing Construction Workforce. In M. B. Jelodar, R. Lovreglio, Z. Feng, D. Paes, V. Kumar, & C. Siriwardana (Eds.), *Creating Capacity and Capability: Embracing Advanced Technologies and Innovations for Sustainable Future in Building Education and Practice* (Vol. 564, pp. 41–52). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-2802-5_4

Wang, X., & Ren, Y. (2026). The Spillover Effects of E-Commerce Platform Algorithmic Governance: A Focus on Ride-Hailing Drivers' High-Calorie Food Consumption. *Journal of Theoretical and Applied Electronic Commerce Research*, *21*(2), 66. https://doi.org/10.3390/jtaer21020066

Wu, D., Xie, Y., & Pei, Y. (2025). Moving from Unethical Competition to Sustainability: Understanding the Impact of Competitive Ethics Review on Carbon Reduction of Digital Enterprise. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-025-06135-1