

Human in the Loop: Understanding Employee Perception in AI-Based Performance Management System

Prof.Dr. Mahmut Arslan

Prof. Dr. , Ibn Haldun University, Istanbul, Turkey
mahmut.arslan@ihu.edu.tr, (Orcid: 0000-0002-5772-3627)

Dr. Alireza Ghezel*

PhD, Researcher, Ibn Haldun University, Istanbul, Turkey
Alirezaghezel@ihu.edu.tr (Orcid: 0009-0009-4436-5718)

Nour El Hoda Tarabah

PhD C. ,Ibn Haldun University,Istanbul, Turkey
Nour.tarabah@stu.ihu.edu.tr, (Orcid: [0000-0001-7864-2394](https://orcid.org/0000-0001-7864-2394))

Muhammed Ali Selçuk

Master Student, Ibn Haldun University, Istanbul, Turkey
muhammed.selcuk@stu.ihu.edu.tr (Orcid: 0009-0005-0459-7638)

Noor ul Saba

Master Student, Ibn Haldun University, Istanbul, Turkey
noor.ulsaba@stu.ihu.edu.tr (Orcid: 0000-0001-9256-2841)

Keywords: AI-based performance management system; algorithmic management; human-in-the-loop; task-contingent legitimacy; employee perceptions; human-AI collaboration

Abstract

As organizations increasingly deploy algorithmic management systems to support performance evaluation, employee perceptions of AI-based performance management systems remain underexplored, particularly in knowledge-intensive environments where performance cannot be reduced to standardized metrics alone. This study examines how employees perceived and experienced an internally developed AI-based performance management system structured around the planning, monitoring, and review stages of the performance management cycle. Using a qualitative case study design, 19 semi-structured interviews were analyzed through Braun and Clarke's reflexive thematic analysis, supported by MAXQDA. Four themes were identified: emotionally-disconnected efficiency, task-bound trust, human-in-the-loop safeguard, and algorithmic fairness tension. Employees valued AI for speed, objectivity, documentation, and reduced managerial favoritism, but raised concerns about impersonal feedback, surveillance-like monitoring, contextual blindness, hidden bias, and opacity. Extending prior task-dependent algorithm-trust research, the findings show that AI-PMS legitimacy is not shaped by task type alone, but by the combined configuration of task type, performance-management stage, system transparency, and human oversight. AI was perceived as more legitimate for quantifiable KPI-based evaluation, but less legitimate for qualitative, relational, developmental, and career-consequential judgments. The study contributes to algorithmic management and AI-HRM research by moving beyond binary acceptance or resistance models toward a conditional legitimacy framework. For managers responsible for developing and implementing AI-based performance management systems, the findings highlight the importance of task-calibrated AI deployment, human-in-the-loop governance, transparent evaluation logic, and empathetic feedback practices to support the effective use of AI in critical HR functions such as performance management.

Introduction

Artificial intelligence (AI) increasingly shapes managerial decision-making in knowledge-intensive organizations. AI-enabled systems now support hiring, task allocation, workforce analytics, performance monitoring, and promotion decisions across many organizations (Fenwick et al., 2024). Within this broader shift toward algorithmic management, performance management systems (PMS) have emerged as one of the most strategically sensitive domains of AI adoption. This development is also relevant to technology-management contexts because managers responsible for AI-enabled systems must decide how algorithmic outputs should support, but not replace, judgment in evaluating complex knowledge work (Black & van Esch, 2020).

Traditionally, PMS relied on periodic appraisals, managerial judgment, and supervisor feedback (Carroll & Schneier, 1982; Walsh, 2003). More recently, organizations have introduced AI-enabled technologies, including machine learning algorithms, predictive analytics, automated dashboards, and feedback tools, to improve the speed, consistency, and analytical capacity of evaluation processes (Varma et al., 2024). Existing research suggests that AI-based HR systems may reduce some forms of human bias, support continuous feedback, and improve alignment between employee performance and organizational objectives (Ekuma, 2024; Madanchian et al., 2023; Fenwick et al., 2024). However, the integration of AI into PMS also raises broader concerns associated with algorithmic management.

Algorithmic management refers to the use of algorithmic systems to monitor, evaluate, coordinate, and control work through automated or semi-automated processes (Kellogg et al., 2020). Prior research shows that algorithms are not neutral technical instruments. They can reshape managerial authority, workplace control, employee autonomy, and perceptions of fairness. Parent-Rocheleau and Parker (2022) further argue that algorithmic systems influence job demands and job resources through monitoring, goal setting, performance management, scheduling, compensation, and termination practices. These concerns are especially relevant in AI-based PMS because performance evaluation represents one of the most visible forms of algorithmic control within organizations.

Although AI-supported PMS may improve consistency and reduce overt favoritism, algorithmic systems may also reproduce hidden bias, intensify surveillance concerns, and generate perceptions of injustice when employees cannot understand how evaluative decisions are produced (Bankins et al., 2022; Köchling & Wehner, 2020; Langer & Landers, 2021; Majrashi, 2025). Langer and Landers (2021) show that employees evaluate AI-supported decisions through perceptions of fairness, trustworthiness, transparency, and system design. Bankins et al. (2022) similarly argue that AI-supported HR decisions may weaken perceptions of dignity, interactional justice, and empathy when employees experience evaluative systems as impersonal or dehumanizing. These findings suggest that the legitimacy of AI-based PMS depends on more than technical efficiency or analytical accuracy.

This creates the central research problem addressed in the present study. Organizations increasingly adopt AI-based PMS to improve evaluation efficiency and objectivity, yet limited research examines how employees interpret, trust, resist, or accept these systems in practice.

Existing research has largely focused on AI adoption, technical functionality, and organizational outcomes (Palos-Sánchez et al., 2022; Madanchian et al., 2023), while comparatively fewer studies examine employees as active sense-makers who interpret algorithmic evaluation through emotional, ethical, relational, and career-related concerns. This limitation is particularly important in knowledge-intensive environments where performance cannot be fully reduced to numerical indicators alone.

Prior research has already shown that trust in algorithms is task-dependent. Algorithm aversion research indicates that individuals may resist algorithmic judgments after observing errors, especially when decisions affect personally meaningful outcomes (Dietvorst et al., 2015). Algorithm appreciation research, in contrast, shows that individuals may prefer algorithmic judgment when systems appear objective, consistent, and analytically reliable (Logg et al., 2019). More directly, Castelo et al. (2019) demonstrate that people are more willing to rely on algorithms for objective and measurable tasks than for subjective tasks, while Lee (2018) shows that algorithmic management decisions are evaluated differently depending on task characteristics, with implications for fairness, trust, and emotion. These studies establish task-dependence as an important feature of algorithmic trust. However, AI-based performance management raises an additional question: how does task-dependence operate across a complete performance-management cycle in which planning, monitoring, and review involve different levels of measurement, transparency, and human interpretive responsibility?

This study addresses these gaps by examining employee perceptions of AI-based PMS within a university-based knowledge-intensive setting. Such settings are useful for studying AI-supported performance evaluation because work outcomes combine measurable outputs with qualitative, relational, developmental, and professional contributions. The analytical transferability of the study lies in this mixed evaluation logic. Similar challenges are faced by managers in technology-management contexts, including AI product teams, software development units, and technology-intensive organizations, where performance assessment often involves measurable indicators as well as tacit knowledge, collaboration, innovation, and professional judgment (Huang, 2009; Edmondson et al., 2003). The research question guiding this study is: How do employees perceive and experience AI-based performance management systems in their organization?

This study develops a qualitative understanding of how employees interpret AI-based PMS in relation to trust, fairness, emotional experience, task appropriateness, transparency, and human oversight. The contribution is not to claim task-dependent algorithm trust as a new phenomenon. Rather, the study extends prior task-dependent algorithm-trust research by showing how AI-PMS legitimacy is shaped by the combined configuration of four conditions: the type of evaluative task, the stage of the performance-management cycle, the transparency of the system's evaluative logic, and the presence of meaningful human oversight. This framing helps explain why the same AI-PMS may be perceived as legitimate for structured planning, KPI monitoring, and analytical reporting, yet less legitimate for developmental feedback, qualitative interpretation, and career-consequential judgments. The findings also provide practical implications for managers

responsible for implementing AI-supported evaluation systems in knowledge-intensive work settings.

Literature Review

Algorithmic Management and AI-Based Performance Management Systems

The growing use of artificial intelligence in organizations has expanded managerial reliance on algorithmic systems for monitoring, evaluating, and coordinating work. In human resource management, AI-enabled performance management systems (PMS) increasingly support functions such as productivity tracking, workforce analytics, feedback generation, and performance evaluation (Varma et al., 2024). Organizations adopt these systems to improve analytical capacity, accelerate decision-making, and increase consistency across evaluation processes (Ekuma, 2024; Fenwick et al., 2024). These developments are particularly visible in technology-intensive and knowledge-based organizations where managers seek measurable indicators for project delivery, technical contribution, innovation outcomes, and team effectiveness (Amaya Silva & Holweg, 2024; Huang, 2009).

The integration of AI into PMS reflects a broader transition toward algorithmic management. Kellogg et al. (2020) define algorithmic management as the use of algorithmic systems to structure, monitor, optimize, and control work through automated or semi-automated mechanisms. Parent-Rocheleau and Parker (2022) further show that such systems reshape job demands and resources through intensified monitoring and altered perceptions of autonomy and organizational support. Complementing this control perspective, Jarrahi (2018) argues that AI creates the greatest value when its analytical capacity augments, rather than replaces, human intuition and contextual judgment in decisions characterized by uncertainty and equivocality.

Existing research identifies several operational advantages associated with AI-based PMS. AI-supported systems can process large datasets rapidly, identify performance patterns, generate real-time feedback, and reduce inconsistency in evaluation practices (Madanchian et al., 2023; Varma et al., 2024). However, recent research indicates that AI is particularly effective in processing codified, high-volume, and specialized knowledge, whereas tacit knowledge, cross-context integration, and complex nonroutine judgment remain primarily human contributions (He & Burger-Helmchen, 2025). Evidence from AI-augmented design thinking similarly shows that AI can accelerate analytical and creative tasks, but poorly structured involvement may weaken collaboration, ownership, and process flow, making the design of human and AI roles consequential for effective augmentation (Polster et al., 2025).

Köchling and Wehner (2020), in their systematic review of algorithmic decision-making in HRM, demonstrate that AI systems may reproduce historical bias embedded in organizational data and evaluation criteria. Algorithms can therefore reinforce discriminatory outcomes even when organizations implement them to reduce subjectivity and improve objectivity. Similar concerns emerge in discussions of transparency and explainability. Employees often struggle to understand how algorithmic systems generate evaluations or prioritize performance indicators, particularly

when organizations use opaque machine learning models (Langer & Landers, 2021). Limited transparency may weaken trust in the evaluation process and intensify concerns regarding accountability and procedural fairness.

These concerns become especially significant in performance management because evaluation systems shape compensation, promotion opportunities, developmental feedback, and professional legitimacy. Bankins et al. (2022) argue that AI-supported HR decisions may generate perceptions of dehumanization, emotional detachment, and reduced interactional justice when employees perceive evaluative systems as overly mechanistic or context-insensitive. Noponen et al. (2024) similarly suggest that employees interpret algorithmic managerial systems through broader perceptions of legitimacy, organizational control, and fairness rather than through technical functionality alone. Recent research on generative AI adoption also shows that organizational integration depends on more than technical capability: data privacy, system compatibility, employee resistance, training, and governance influence successful implementation, while SMEs and large organizations face different resource and scaling conditions (Al-Kfairy, 2025). Taken together, the literature presents a central tension: organizations adopt AI-enabled PMS for efficiency and analytical capability, yet employees may question its legitimacy when systems appear opaque, inflexible, or disconnected from the human dimensions of work.

Employee Perceptions of Algorithmic Evaluation

Employee perceptions play a central role in determining the effectiveness and legitimacy of performance management systems because evaluation influences compensation, promotion opportunities, professional reputation, developmental feedback, and employees' broader relationship with the organization. Organizational justice research suggests that employee reactions to evaluation systems are shaped by procedural fairness, interactional fairness, and perceptions of managerial legitimacy (Gilliland, 1993; Colquitt et al., 2001). In AI-based PMS, these concerns become more complex because algorithmic systems increasingly participate in evaluative processes traditionally associated with human judgment. Trust is particularly central in AI-supported evaluation because it guides whether employees rely on algorithmic outputs. Foundational trust-in-automation research conceptualizes appropriate reliance as a matter of calibration and functional specificity: trust should correspond to the system's actual capability in the particular task and context in which it is used (Lee & See, 2004). More recent research similarly identifies system performance, explainability, task characteristics, and the perceived role of AI as important determinants of trust (Glikson & Woolley, 2020).

Research on AI-supported HR decisions shows that employees often hold competing perceptions regarding algorithmic evaluation. On one hand, AI-supported systems may appear more consistent, standardized, and less vulnerable to favoritism than human evaluators (Logg et al., 2019). Organizations therefore frequently justify AI adoption through claims of neutrality, efficiency, and evidence-based decision-making (Köchling & Wehner, 2020). On the other hand, employees may question whether algorithmic systems can fully account for contextual complexity, individual circumstances, or qualitative dimensions of performance. Langer and Landers (2021)

argue that perceptions of fairness in AI-supported decisions depend heavily on transparency, explainability, and employees' understanding of how algorithmic evaluations are produced.

Concerns regarding emotional legitimacy further complicate employee responses to AI-based evaluation systems. Performance management involves communication, recognition, developmental guidance, and interpersonal interaction in addition to measurement and monitoring (DeNisi & Murphy, 2017). Bankins et al. (2022) argue that AI-supported HR decisions may reduce perceptions of dignity, empathy, and individualized treatment when employees experience algorithmic systems as impersonal or emotionally detached. Similar concerns appear in research examining digitalized forms of managerial control, where employees report discomfort when algorithmic systems replace relational or context-sensitive managerial interaction (Parent-Rocheleau & Parker, 2022). These tensions may become more salient in knowledge-intensive environments where employees expect evaluation processes to incorporate contextual understanding, developmental dialogue, and professional judgment.

Existing research also suggests that trust in algorithmic systems is conditional rather than absolute. Studies on algorithm aversion demonstrate that individuals may reject algorithmic judgments after observing errors, especially when decisions affect personally meaningful outcomes such as professional evaluation or career progression (Dietvorst et al., 2015). However, resistance to algorithms is not uniform. Dietvorst et al. (2018) show that allowing individuals even limited control over algorithmic recommendations can reduce algorithm aversion, while Logg et al. (2019) demonstrate that individuals may prefer algorithmic judgments when systems appear objective and analytically reliable. Castelo et al. (2019) further show that people are more willing to rely on algorithms for objective tasks than for subjective tasks. Similarly, Lee (2018), in the context of algorithmic management, shows that task characteristics shape perceptions of fairness, trust, and emotion in response to algorithmic decisions. These studies provide a strong foundation for understanding task-dependent trust, but they leave open how such task-dependence unfolds within performance management systems that combine planning, monitoring, review, feedback, and developmental judgment.

These findings indicate that employee acceptance of AI-based PMS depends on more than technical system performance. Employees appear to evaluate algorithmic systems according to whether AI is perceived as appropriate for specific evaluative tasks and whether meaningful human oversight remains present within the decision-making process. Current literature therefore explains important dimensions of fairness, transparency, trust, and algorithmic evaluation, yet provides more limited understanding of how employees interpret and negotiate these competing perceptions within everyday organizational settings.

Conceptual Positioning of the Present Study

Existing research provides important insight into the operational and ethical implications of AI-based performance management systems. Prior studies explain how AI-supported evaluation may improve consistency, increase analytical capability, and support data-driven decision-making (Ekuma, 2024; Varma et al., 2024). Other studies emphasize concerns surrounding transparency,

fairness, algorithmic bias, emotional detachment, and managerial legitimacy in AI-supported HR decisions (Köchling & Wehner, 2020; Bankins et al., 2022; Langer & Landers, 2021). Research on algorithm aversion and algorithm appreciation further demonstrates that employee responses to AI-supported evaluation are often conditional and context dependent rather than uniformly positive or negative (Dietvorst et al., 2015; Logg et al., 2019).

Despite these advances, comparatively fewer studies examine how employees experience AI-based performance evaluation as part of everyday organizational life, particularly in knowledge-intensive environments where performance cannot be fully reduced to standardized quantitative indicators. In universities and similar professional settings, performance evaluation frequently involves qualitative contribution, collaboration, creativity, developmental interaction, and professional discretion. Similar tensions increasingly emerge in engineering, consulting, and other expertise-driven environments where organizations seek data-driven evaluation while simultaneously managing complex and highly contextual forms of work.

The present study addresses these gaps by adopting a qualitative interpretivist approach to examine how employees perceive and experience AI-based performance management within a university context. Rather than focusing primarily on system adoption or technical performance, the study explores how employees interpret AI-supported evaluation in relation to trust, fairness, legitimacy, emotional experience, transparency, and human oversight. Building on prior research on task-dependent algorithm trust and algorithmic management (Castelo et al., 2019; Lee, 2018), the study examines how task-dependence operates within the specific structure of the performance-management cycle. It therefore positions AI-PMS legitimacy as a configuration of task type, cycle stage, system transparency, and human interpretive authority, rather than as a response to task type alone.

Research Design and Context

This study adopted an interpretivist qualitative approach to explore how employees perceive and experience AI-based performance management systems within their organizational context. An interpretivist perspective was appropriate because the study sought to understand how employees subjectively interpret AI-supported evaluation processes and organizational practices rather than test predefined causal relationships. The research was conducted as a qualitative case study within one of Türkiye's universities, where employees had direct experience using an AI-based performance management system (AI-PMS). The university represented a knowledge-intensive organizational setting in which performance evaluation combined both quantitative and qualitative dimensions of work. Data were collected through semi-structured interviews and analyzed using Braun and Clarke's (2006, 2022) reflexive thematic analysis approach.

The AI-PMS examined in this study was an internally developed platform designed through collaboration between the university's Human Resources Department and internal software development specialists. Technically, the system combined structured performance-data tracking, rule-based workflow automation, dashboard-based analytics, and AI-supported generation of performance summaries, feedback prompts, and developmental suggestions. Several participants,

particularly those in HR and system-related roles, had experience not only as users of the system but also as implementation stakeholders involved in refining or improving the AI-PMS in collaboration with technical staff. The system had been in active use for one full academic performance cycle at the time of data collection. The platform was structured around Armstrong and Taylor's (2023) performance management cycle, which conceptualizes performance management as an ongoing process consisting of planning, monitoring, and review stages. During the planning stage, the AI assistant supported managers and employees by analyzing prior performance data and assisting in the formulation of SMART performance goals. During monitoring, the system supported documentation of critical incidents, ongoing performance tracking, and real-time monitoring activities. During the review and feedback stage, the AI assistant helped managers prepare timely feedback, generate developmental suggestions, and support reward-related decisions based on performance-related indicators.

The interview protocol was designed in alignment with both the performance management process and the broader literature on employee perceptions of AI-supported evaluation systems. Questions explored employee perceptions related to trust, fairness, transparency, emotional experience, perceived usefulness, perceived ease of use, acceptance, human oversight, and developmental value. These dimensions were selected because prior organizational behavior and AI-HRM research has identified them as central factors shaping employee reactions toward AI-supported managerial decisions (Bankins et al., 2022; Davis, 1989; DeNisi & Murphy, 2017; Glikson & Woolley, 2020; Langer & Landers, 2021). The protocol also examined employee experiences across the planning, monitoring, and review stages of the performance management cycle to ensure participants reflected on concrete interactions with the AI-PMS. The interview protocol was reviewed by the research team prior to data collection to ensure clarity and alignment with the study's interpretive focus. The full interview protocol is available from the corresponding author upon request.

Data Collection and Analysis

Participants were selected through purposive sampling. The primary inclusion criterion was sufficient familiarity with the university's performance management process and practical experience using the AI-PMS platform. The sample included both academic and administrative staff members to capture diverse organizational perspectives regarding AI-supported evaluation practices. Invitations were distributed through internal organizational communication channels to employees who met the inclusion criteria, and participation was fully voluntary. Purposive sampling was selected because the study sought information-rich participants capable of providing detailed reflections on AI-supported performance management rather than statistically representative responses. Attention was also given to variation in organizational role and institutional experience to limit the risk of relying only on particularly vocal supporters or critics of the system.

A total of 19 semi-structured interviews were conducted, ranging approximately between 40 and 75 minutes. Data collection and analysis occurred iteratively throughout the research process.

The study considered both code saturation and meaning saturation in determining sampling sufficiency. Following Hennink et al. (2017), repetitive coding patterns began to stabilize after approximately 14 interviews, indicating code saturation. Additional interviews were conducted to deepen interpretive understanding, refine conceptual distinctions, and strengthen meaning saturation, which is particularly important in reflexive thematic analysis where interpretive depth is prioritized alongside code recurrence. Data collection concluded after 19 interviews when additional interviews no longer generated substantively new conceptual insights relevant to the research focus.

Interview data were analyzed using Braun and Clarke's (2006, 2022) reflexive thematic analysis approach. To enhance procedural transparency, the analysis was also informed by the step-by-step thematic analysis framework proposed by Naeem et al. (2023). The analysis began with repeated reading of interview transcripts to achieve familiarity with the data and identify preliminary interpretive patterns. Initial codes were generated across the dataset and iteratively refined through comparison between transcripts, codes, and emerging thematic interpretations. Related codes were then grouped into candidate themes, which were reviewed and refined in relation to both the dataset and the research question. The analysis moved between semantic meanings expressed directly by participants and latent meanings related to underlying assumptions, tensions, and interpretations. MAXQDA supported data organization, coding management, and retrieval throughout the analytical process. The complete six-phase analytical walkthrough, including illustrative codes, theme development decisions, and supporting participant quotations, is provided in Appendix A. Given the interpretive nature of reflexive thematic analysis, researcher positionality was treated as an integral component of the analytical process rather than a source of bias to be eliminated. The research team's professional background in organizational behavior, human resource management, and coaching supported contextual understanding during data interpretation. At the same time, reflexive attention was maintained throughout coding and theme development to reduce the imposition of prior assumptions on participants' narratives. Reflexive memos were written following interviews and throughout the analytical process to document emerging interpretations, analytical tensions, and evolving understandings of the data. Iterative engagement with the data and continuous comparison across participant accounts supported analytical transparency and reflexive accountability throughout the study (Braun & Clarke, 2022).

The study addressed established qualitative trustworthiness criteria, including credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985). Credibility was supported through iterative coding, semantic and latent coding strategies, constant comparison across all transcripts, and inclusion of participants from diverse academic and administrative roles. Transferability was strengthened through contextual description of the organizational setting, the structured interview protocol, and the participant profile presented in Table 1. Dependability was addressed through transparent documentation of analytical reasoning and theme development across all phases of the thematic analysis. Confirmability was supported through reflexive memoing, iterative comparison between data and interpretations, and systematic grounding of

analytical claims in participant narratives, consistent with Braun and Clarke's (2022) reflexive epistemological position. Ethical considerations were observed throughout the study. Participants provided informed consent prior to interviews, participation was voluntary, and all interview data were anonymized to protect confidentiality and participant privacy.

Findings

Descriptive Profile of Participants

The descriptive profile of the 19 participants reflects variation in gender, age, organizational role, and institutional experience (Table 1). The sample was nearly gender-balanced, with 10 females and 9 males. Most participants were in the 30–39 age range ($n = 8$), followed by participants aged 20–29 ($n = 5$), with a smaller number in older age brackets. In terms of role profile, participants included academic and research staff, HR and system implementation stakeholders, and senior administrative users. This role composition was appropriate for the study because the AI-PMS was experienced both as a performance evaluation tool and as an internally implemented organizational system. The analysis does not aim to compare occupational groups. Rather, it focuses on shared and contrasting cross-role perceptions of AI-supported performance evaluation within a knowledge-intensive organizational setting.

Table 1. Descriptive Profile of Participants

Variable	Category	n
Gender	Female (10); Male (9)	19
Age Range	20-29 (5); 30-39 (8); 40-49 (4); 50-59 (2)	19
Position	Academic and research staff (13); HR/system implementation stakeholders (4); Senior administrative users (2)	19
Experience at University	≤ 2 years (7); 3-5 years (6); > 5 years (6)	19
Total Experience	≤ 5 years (6); 6-10 years (7); > 10 years (6)	19

Thematic Findings

The reflexive thematic analysis identified four interrelated themes explaining how employees perceived and experienced the AI-PMS: emotionally-disconnected efficiency, task-bound trust, human-in-the-loop safeguard, and algorithmic fairness tension. The detailed analytical walkthrough, including the initial code families, preliminary thematic structure, final theme definitions, and illustrative high-frequency codes, is provided in Appendix A. In the main text, the findings are organized around the four final themes and the process framework derived from the thematic synthesis.

Theme 1: Emotionally-Disconnected Efficiency

Participants valued the AI-PMS for improving speed, consistency, documentation, and objectivity. At the same time, they described AI-supported evaluation as emotionally distant and

limited in its ability to understand the human side of performance. This theme captures the central tension between operational efficiency and relational disconnection.

One participant described this tension as follows:

“The system is definitely faster and more organized than traditional evaluation methods. Everything is documented, and managers cannot easily ignore performance data. But at the same time, it feels emotionally distant. Sometimes it feels like the system sees employees as numbers rather than people.” (I9)

Another participant emphasized that AI could support continuous and objective monitoring, but could not fully interpret stress, motivation, or personal circumstances:

“AI can monitor performance continuously and objectively, but it cannot understand the human side of work. Stress, personal difficulties, motivation, or emotional pressure cannot always be translated into data.” (I4)

Theme 2: Task-Bound Trust

Trust in the AI-PMS was not general or unconditional. Participants expressed stronger trust when the system was used for quantitative, structured, and KPI-based tasks, such as tracking indicators, organizing performance evidence, or producing analytical summaries. Trust weakened when the system was perceived as evaluating qualitative, relational, developmental, or context-sensitive aspects of performance.

As one participant explained:

“For numerical evaluations or tracking performance indicators, AI is probably more reliable than humans because it processes data consistently. But when performance evaluation requires interpretation, mentoring, or understanding people, I still trust human judgment more.” (I15)

This theme shows that employees did not simply accept or reject AI-based evaluation. Instead, they differentiated between tasks where AI appeared analytically useful and tasks where human interpretation remained necessary.

Theme 3: Human-in-the-Loop Safeguard

Participants repeatedly emphasized that AI should support, but not replace, managerial judgment in performance evaluation. They preferred hybrid arrangements in which AI generated evidence, summaries, and recommendations, while human managers retained responsibility for contextual interpretation, feedback delivery, and final evaluative judgment.

One participant stated:

“AI can support managers, but it should not become the final decision-maker. Employees need to know that a human being still reviews the process and understands the broader context behind the evaluation.” (I7)

This theme reflects the importance of visible human oversight. Employees were more willing to accept AI-supported evaluation when they believed that managers could review, contextualize, question, or override AI-generated outputs.

Theme 4: Algorithmic Fairness Tension

Participants held mixed perceptions of algorithmic fairness. Some viewed the AI-PMS as potentially more objective than traditional manager-centered evaluation because it could reduce favoritism and increase consistency. Others argued that AI could reproduce hidden bias if the data, assumptions, indicators, or system design were biased.

As one participant explained:

“The system may reduce personal favoritism, but algorithms are still designed by people. If the data or assumptions behind the system are biased, the results may also become unfair in a different way.” (I11)

This theme shows that employees did not treat fairness as a purely technical outcome. Fairness was evaluated through the interaction of perceived objectivity, transparency, contextual fit, and the possibility of human review.

Task-Contingent AI-PMS Legitimacy Framework

Taken together, the four themes suggest that employee perceptions of AI-PMS legitimacy are shaped by the combined configuration of task type, performance-management stage, system transparency, and human oversight. AI was perceived as more legitimate for structured planning, KPI monitoring, documentation, and analytical reporting, but less legitimate for developmental feedback, qualitative interpretation, and career-consequential judgments. Figure 1 synthesizes this interpretation as a process framework across the planning, monitoring, and review stages of the performance management cycle.

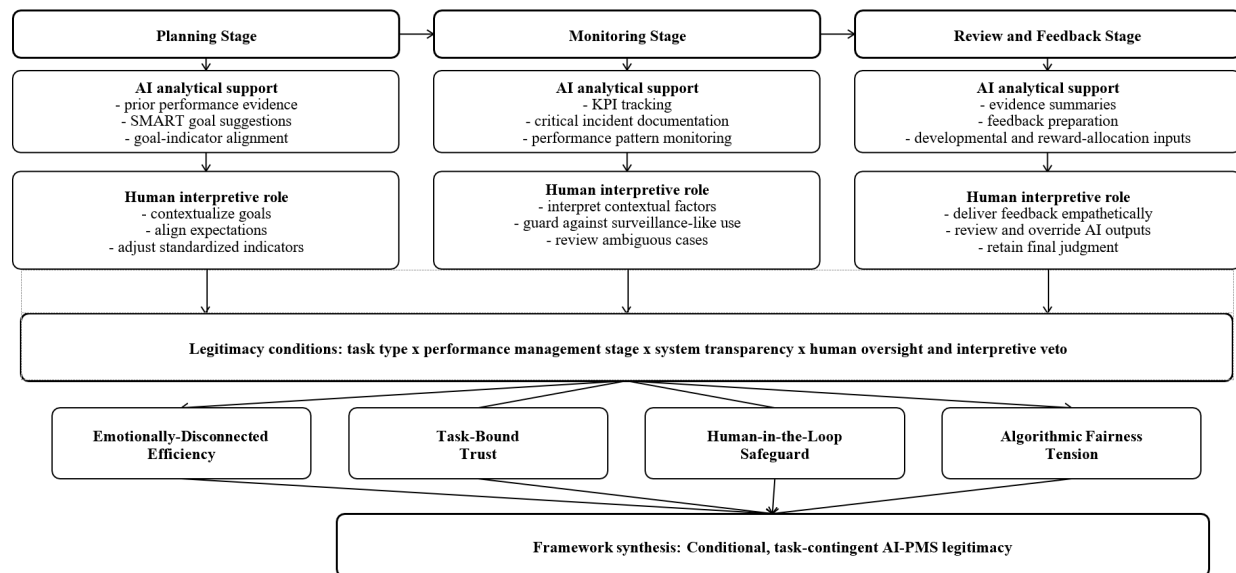


Figure 1. A Process Framework of Task-Contingent AI-PMS Legitimacy Across the Performance Management Cycle

Discussion

This study contributes to the growing literature on AI-supported performance management by showing that employee perceptions of AI-based PMS are multidimensional, negotiated, and shaped by the relational and contextual characteristics of evaluative work. Participants valued the operational strengths of AI-supported evaluation, including speed, consistency, documentation quality, and analytical support, while also expressing concerns regarding emotional detachment, contextual insensitivity, opacity, and overreliance on algorithmic outputs. These findings suggest that employee responses to AI-based PMS cannot be understood solely through technology acceptance or efficiency-oriented perspectives. Instead, employees evaluate AI-supported performance management through a broader legitimacy logic that combines instrumental, procedural, relational, and ethical considerations, consistent with prior research on algorithmic management, AI-supported HR decisions, and organizational justice (Bankins et al., 2022; Kellogg et al., 2020; Langer & Landers, 2021; Parent-Rochelleau & Parker, 2022).

The central theoretical contribution of this study is to extend task-dependent algorithm-trust research into the context of AI-based performance management. Prior research has shown that people are generally more willing to rely on algorithms for objective, measurable, and structured tasks than for subjective, emotional, or human-centered judgments (Castelo et al., 2019; Lee, 2018). The present findings support this task-dependent logic, but also show that task type alone does not fully explain employee perceptions of AI-PMS legitimacy. In this study, legitimacy was shaped by a configuration of four conditions: whether the evaluative task was measurable or interpretive, whether the AI-PMS was operating in the planning, monitoring, or review stage, whether the system's evaluative logic was transparent or opaque, and whether accountable human oversight remained available. This combined configuration explains why the same AI-PMS could be viewed as useful and legitimate for SMART goal formulation, KPI tracking, documentation, and analytical reporting, yet less legitimate for developmental feedback, qualitative interpretation, and career-consequential judgments.

The findings also extend prior research on AI-supported HRM and algorithmic management by showing how efficiency, emotional legitimacy, fairness, transparency, and human oversight operate together within a single evaluative process. Existing literature has documented these concerns in relation to AI-enabled organizational systems, automation, algorithmic fairness, and human oversight (Chun et al., 2024; Fenwick et al., 2024; Langer & Landers, 2021; Majrashi, 2025; Shulner-Tal et al., 2022). The present findings deepen this literature by showing that trust in AI-supported evaluation is not only task-calibrated but also stage-sensitive and governance-dependent. Employees expressed stronger confidence in AI-supported functions related to goal structuring, KPI tracking, documentation, monitoring, and analytical reporting, but remained skeptical toward AI involvement in developmental feedback, mentoring, motivation, qualitative assessment, and final evaluative judgment. This pattern suggests that AI-PMS should be understood as a staged and governed human-AI evaluation process rather than as a uniform technological intervention.

The findings further contribute to debates on algorithmic fairness and transparency. Several participants perceived AI-based PMS as more objective and less vulnerable to favoritism than traditional manager-centered evaluation. This supports prior studies suggesting that algorithmic systems may increase consistency and reduce some forms of interpersonal bias in HR decision-making (Bankins et al., 2022; Köchling & Wehner, 2020). At the same time, participants expressed concerns regarding hidden algorithmic bias, limited explainability, and uncertainty surrounding how evaluative outputs were generated. Employees therefore did not interpret fairness as a purely technical property of the system. Instead, fairness perceptions were shaped by transparency, contextual appropriateness, and the continued presence of meaningful human judgment. This finding aligns with research showing that employees evaluate AI-supported decisions not only according to outcome consistency but also according to procedural legitimacy, explainability, and interpretive transparency (Langer & Landers, 2021; Shulner-Tal et al., 2022).

A further contribution of this study concerns the emotional and relational dimensions of AI-supported performance management. Participants repeatedly described AI-generated feedback as emotionally distant, impersonal, or insufficiently responsive to the human aspects of work. Prior studies have similarly noted concerns regarding dehumanization, dignity, respectful treatment, and interactional justice in AI-supported HR practices (Bankins et al., 2022; Bankins et al., 2024; Chun et al., 2024). The present findings show how these concerns become particularly salient in performance management contexts where evaluation directly influences employee identity, motivation, development, and psychological safety. This represents an important contextual boundary condition within the AI-HRM literature. AI-supported evaluation may be perceived as legitimate in structured and quantifiable tasks, but less legitimate in contexts involving interpersonal judgment, developmental interaction, qualitative interpretation, and career-related consequences (Castelo et al., 2019; Glikson & Woolley, 2020).

The empirical contribution of this study is grounded in a university-based knowledge-intensive setting, but its analytical transferability is relevant for technology-management contexts in which managers evaluate complex work through both measurable indicators and professional judgment. AI product teams, software development units, and technology-intensive organizations often assess performance through codified outputs such as milestones, documentation quality, defect rates, delivery timelines, and KPI progress. At the same time, these settings also depend on less easily codified contributions, including tacit knowledge, collaboration, mentoring, problem-solving complexity, innovation potential, and contextual performance (Amaya Silva & Holweg, 2024; Edmondson et al., 2003; Huang, 2009). These similarities support the analytical transferability of the framework to technology-management settings, without treating such settings as part of the empirical sample. This interpretation is consistent with algorithmic management research emphasizing human accountability, interpretive flexibility, and governance in AI-supported organizational systems (Kellogg et al., 2020; Parent-Rocheleau & Parker, 2022).

Practical Implications

The findings offer practical guidance for managers responsible for developing, implementing, or governing AI-enabled performance evaluation systems. The central implication is that AI-PMS should be designed as a task-calibrated decision-support system rather than as a universal evaluator. Employees in this study did not reject AI-supported evaluation as a whole; rather, they differentiated between tasks where AI appeared useful and tasks where human judgment remained necessary. Managers should therefore begin by mapping performance-management activities according to their level of measurability, contextual complexity, and career consequence.

First, AI-PMS is most appropriate for structured and evidence-based activities such as goal documentation, KPI tracking, performance-pattern monitoring, evidence summaries, workflow reminders, and analytical reporting. These activities are relatively codified and allow AI to support consistency, documentation, and visibility. However, AI should be used more cautiously in tasks requiring developmental feedback, mentoring, qualitative interpretation, motivational support, or career-related judgment. This recommendation is consistent with the study's finding that employees trusted AI more for quantifiable and KPI-based tasks but placed greater value on human judgment when evaluation involved interpretation, relational understanding, or contextual sensitivity. It also aligns with research on calibrated reliance, which suggests that trust should correspond to the actual capability of the system in a specific task and context (Lee & See, 2004).

Second, managers should make human oversight visible rather than treating it as an implicit safeguard. In this study, employees were more comfortable with AI-supported evaluation when they believed that managers could review, contextualize, question, or override AI-generated outputs. Organizations should therefore clarify who reviews AI-generated recommendations, which decisions remain human-led, when managerial override is expected, and how employees can request clarification or correction. This does not mean that managers should ignore AI outputs. Rather, AI should provide structured evidence and analytical support, while managers remain accountable for interpretation, feedback delivery, and final evaluative judgment. This reflects the broader logic of human-AI augmentation, where AI's analytical capacity complements rather than replaces human contextual judgment (Jarrahi, 2018).

Third, transparency should be treated as a managerial practice, not only as a technical system feature. Employees need to understand what data the AI-PMS uses, what performance indicators are included, how outputs are generated, and how those outputs influence evaluation, development, or reward decisions. Managers do not need to explain every technical detail of the system, but they should provide a clear and usable explanation of the evaluation logic. The findings suggest that opacity can weaken perceived fairness even when the system is intended to improve objectivity. Therefore, organizations should develop simple communication materials, manager talking points, and employee-facing explanations that clarify the role and limits of AI in the performance-management process.

Fourth, AI-PMS implementation should include manager training focused on interpretation and feedback, not only system operation. Managers need to understand how to read AI-generated outputs, identify possible data or context limitations, and translate system-generated insights into

constructive feedback conversations. This is especially important because participants perceived AI-generated feedback as potentially efficient but emotionally distant. Training should therefore prepare managers to combine evidence-based evaluation with empathy, developmental dialogue, and individualized consideration. In practical terms, managers should be trained to ask whether the AI output reflects the broader work context, whether important qualitative contributions are missing, and whether the feedback can be delivered in a way that preserves dignity and motivation.

Fifth, implementation should be adapted to organizational size and resource capacity. Drawing on broader evidence that generative AI adoption pathways differ according to organizational size and resource conditions (Al-Kfairy, 2025), large organizations can support AI-PMS through formal governance structures, audit trails, bias-monitoring routines, manager training, employee appeal mechanisms, and periodic review of evaluation indicators. SMEs may benefit from a more incremental approach, beginning with lower-risk uses such as documentation, KPI tracking, and evidence summaries before extending AI into more sensitive feedback or development processes. Startups should be especially cautious because roles, goals, and evaluation criteria may still be evolving. In such contexts, AI-PMS should support learning, transparency, and coordination rather than impose premature algorithmic control.

Finally, the findings have specific relevance for technology-management contexts, including AI product teams, software development units, and technology-intensive organizations. In these settings, some outputs are measurable through milestones, delivery timelines, documentation quality, defect rates, or KPI progress. However, performance also depends on tacit and collaborative contributions such as problem-solving complexity, mentoring, innovation potential, and cross-functional coordination. Because recent research shows that AI is stronger in processing codified knowledge than in replacing tacit, contextual, and socially embedded judgment (He & Burger-Helmchen, 2025), managers should avoid converting all performance dimensions into algorithmic metrics. A more credible approach is to use AI-PMS to strengthen evidence quality and analytical consistency while preserving human responsibility for contextual interpretation, developmental judgment, and final accountability.

Limitations and Future Research

Several limitations should be considered when interpreting the findings of this study. First, the research was conducted within a single university-based knowledge-intensive context in Türkiye. Although this setting provided a valuable context for examining AI-supported performance evaluation, organizational culture, institutional structure, and national context may influence employee perceptions of AI-based evaluation systems. In addition, the participant group was not designed to compare engineering, software development, or R&D occupational groups as distinct populations. Future studies could therefore examine AI-PMS in engineering teams, software development units, AI product development contexts, R&D-intensive organizations, and technology firms to further assess the transferability and refinement of the framework.

Second, participation was voluntary, which introduces the possibility of self-selection bias. Employees with particularly positive or negative perceptions toward the AI-PMS may have been

more willing to participate in interviews. In addition, because the AI-based PMS was organizationally endorsed, some participants may have moderated critical comments due to social desirability concerns despite confidentiality assurances.

Third, the study captured employee perceptions during the relatively early stages of AI-PMS implementation. Perceptions regarding trust, fairness, transparency, and usability may evolve over time as employees gain greater familiarity with the system and as organizational practices surrounding AI-supported evaluation mature. Longitudinal research could therefore provide deeper understanding of how employee perceptions develop across different implementation stages.

Finally, the interpretive nature of reflexive thematic analysis means that researcher positionality inevitably influenced the analytical process. Although reflexive memoing, iterative coding, and continuous comparison across participant accounts were used to support analytical rigor, alternative interpretations of the data may remain possible. Future research may benefit from combining qualitative approaches with quantitative or longitudinal designs to examine how employee perceptions of AI-supported performance management influence broader organizational outcomes such as engagement, psychological safety, trust, and performance over time.

Conclusion

This study examined employee perceptions of an AI-based performance management system within a university context and identified four interrelated themes shaping employee experience with AI-supported evaluation: emotionally-disconnected efficiency, task-bound trust, human-in-the-loop safeguard, and algorithmic fairness tension. The findings demonstrate that employees simultaneously value the operational advantages of AI-supported evaluation while remaining concerned about emotional detachment, contextual limitations, and the risks of opaque or insufficiently supervised algorithmic decision-making. The study contributes to AI-supported HRM and algorithmic management research by extending prior task-dependent algorithm-trust findings into the performance-management cycle. Rather than presenting task-contingency as entirely new, the findings show that AI-PMS legitimacy depends on how task type interacts with the stage of performance management, the transparency of AI-generated outputs, and the availability of human oversight. Employees therefore differentiated not only between measurable and qualitative tasks, but also between planning, monitoring, and review activities, and between AI outputs that were explainable and human-reviewed versus opaque or insufficiently contextualized. As organizations increasingly integrate AI into performance management processes, the findings highlight the importance of balancing technological efficiency with transparency, human oversight, and psychological legitimacy. Sustainable implementation of AI-based PMS may therefore depend less on automation alone and more on organizations' ability to design systems that employees perceive as analytically useful, contextually appropriate, and relationally credible.

References

- Al-Kfairy, M. (2025). Strategic integration of generative AI in organizational settings: Applications, challenges, and adoption requirements. *IEEE Engineering Management Review*, 53(6), 80–97. <https://doi.org/10.1109/EMR.2025.3534034>
- Amaya Silva, J. R., & Holweg, M. (2024). Using algorithms to improve knowledge work. *Journal of Operations Management*, 70(3), 482–513. <https://doi.org/10.1002/joom.1296>
- Armstrong, M., & Taylor, S. (2023). *Armstrong's handbook of human resource management practice: A guide to the theory and practice of people management* (16th ed.). Kogan Page.
- 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 a human resource management context. *Information Systems Frontiers*, 24(3), 857–875. <https://doi.org/10.1007/s10796-021-10223-8>
- Bankins, S., Ocampo, A. C. G., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, 45(2), 159–182. <https://doi.org/10.1002/job.2735>
- Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, 63(2), 215–226. <https://doi.org/10.1016/j.bushor.2019.12.001>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage.
- Carroll, S. J., & Schneier, C. E. (1982). *Performance appraisal and review systems: The identification, measurement, and development of performance in organizations*. Scott Foresman.
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809–825. <https://doi.org/10.1177/0022243719851788>
- Chun, J. S., De Cremer, D., Oh, E. J., & Kim, Y. (2024). What algorithmic evaluation fails to deliver: Respectful treatment and individualized consideration. *Scientific Reports*, 14, Article 25996. <https://doi.org/10.1038/s41598-024-76320-1>
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86(3), 425–445. <https://doi.org/10.1037/0021-9010.86.3.425>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, 102(3), 421–433. <https://doi.org/10.1037/apl0000085>

- 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. <https://doi.org/10.1037/xge0000033>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can even slightly modify them. *Management Science*, 64(3), 1155–1170. <https://doi.org/10.1287/mnsc.2016.2643>
- Edmondson, A. C., Winslow, A. B., Bohmer, R. M. J., & Pisano, G. P. (2003). Learning how and learning what: Effects of tacit and codified knowledge on performance improvement following technology adoption. *Decision Sciences*, 34(2), 197–224. <https://doi.org/10.1111/1540-5915.02316>
- Ekuma, K. J. (2024). Artificial intelligence and automation in human resource development: A systematic review. *Human Resource Development Review*, 23(2), 199–229. <https://doi.org/10.1177/15344843231224009>
- Fenwick, A., Molnar, G., & Frangos, P. (2024). Revisiting the role of HR in the age of AI: Bringing humans and machines closer together in the workplace. *Frontiers in Artificial Intelligence*, 6, Article 1272823. <https://doi.org/10.3389/frai.2023.1272823>
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18(4), 694–734. <https://doi.org/10.5465/amr.1993.9402210155>
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>
- He, X., & Burger-Helmchen, T. (2025). Evolving knowledge management: Artificial intelligence and the dynamics of social interactions. *IEEE Engineering Management Review*, 53(5), 215–231. <https://doi.org/10.1109/EMR.2024.3519342>
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research*, 27(4), 591–608. <https://doi.org/10.1177/1049732316665344>
- Huang, C.-C. (2009). Knowledge sharing and group cohesiveness on performance: An empirical study of technology R&D teams in Taiwan. *Technovation*, 29(11), 786–797. <https://doi.org/10.1016/j.technovation.2009.04.003>
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>
- 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. <https://doi.org/10.5465/annals.2018.0174>
- Köchling, 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. <https://doi.org/10.1007/s40685-020-00134-w>

- Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior*, 123, Article 106878. <https://doi.org/10.1016/j.chb.2021.106878>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 1–16. <https://doi.org/10.1177/2053951718756684>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- 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. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- Madanchian, M., Taherdoost, H., & Mohamed, N. (2023). AI based human resource management tools and techniques: A systematic literature review. *Procedia Computer Science*, 229, 367–377. <https://doi.org/10.1016/j.procs.2023.12.039>
- Majrashi, K. (2025). Employees' perceptions of the fairness of AI-based performance prediction features. *Cogent Business & Management*, 12(1), Article 2456111. <https://doi.org/10.1080/23311975.2025.2456111>
- Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods*, 22. <https://doi.org/10.1177/16094069231205789>
- Noponen, N., Feshchenko, P., Auvinen, T., Luoma-aho, V., & Abrahamsson, P. (2024). Taylorism on steroids or enabling autonomy? A systematic review of algorithmic management. *Management Review Quarterly*, 74(3), 1695–1721. <https://doi.org/10.1007/s11301-023-00345-5>
- Palos-Sánchez, P. R., Baena-Luna, P., Badicu, A., & Infante-Moro, J. C. (2022). Artificial intelligence and human resource management: A bibliometric analysis. *Applied Artificial Intelligence*, 36(1), Article 2145631. <https://doi.org/10.1080/08839514.2022.2145631>
- Parent-Rochelleau, X., & Parker, S. K. (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review*, 32(3), Article 100838. <https://doi.org/10.1016/j.hrmr.2021.100838>
- Polster, L., Bilgram, V., & Görtz, S. (2025). AI-augmented design thinking: Potentials, challenges, and mitigation strategies of integrating artificial intelligence in human-centered innovation processes. *IEEE Engineering Management Review*, 53(5), 193–214. <https://doi.org/10.1109/EMR.2024.3512866>
- Shulner-Tal, A., Kuflik, T., & Kliger, D. (2022). Fairness, explainability and in-between: Understanding the impact of different explanation methods on non-expert users'

perceptions of fairness toward an algorithmic system. *Ethics and Information Technology*, 24(1), Article 2. <https://doi.org/10.1007/s10676-022-09623-4>

- Varma, A., Pereira, V., & Patel, P. (2024). Artificial intelligence and performance management. *Organizational Dynamics*, 53(1), Article 101037. <https://doi.org/10.1016/j.orgdyn.2024.101037>
- Walsh, M. B. (2003). Refocusing performance management for high performance. *Compensation & Benefits Review*, 35(5), 20–32. <https://doi.org/10.1177/0886368703256939>

Appendix A. Analytical Process and Coding Summary

Analysis Process

To enhance transparency in theme development, this section summarizes how the six phases of reflexive thematic analysis were applied to the 19 interview transcripts.

Phase 1 – Familiarizing Yourself with the Data

All 19 interviews were transcribed verbatim and repeatedly read while writing analytic memos. Initial observations noted a recurring polarity in participants' accounts: the AI-PMS was often perceived as objective, efficient, and time-saving, while also being associated with concerns about emotional distance, surveillance-like monitoring, limited contextual understanding, and opacity. These initial memos informed the later development of the four themes but were not treated as final evidentiary quotations.

Phase 2 – Initial Coding and Code Development

Using MAXQDA, the analysis generated 423 semantic and latent initial codes across the dataset. These codes captured participants' operational, emotional, relational, technical, and evaluative perceptions of the AI-PMS. Consistent with reflexive thematic analysis, codes were used as interpretive units for identifying recurring meanings, tensions, contradictions, and contextual variations across participant accounts. The initial codes were organized into broader code families to support systematic comparison across participants and to prepare for candidate theme development. The initial code families are presented in Table A1 to show how early codes were organized before candidate themes were developed. An illustrative high-frequency coding summary is provided in Table A4 for additional transparency.

Table A1. Initial Code Families Used in the Coding Process

Initial code family	Illustrative initial codes	Analytical focus at this stage
Operational perceptions	Efficiency/time-saving; objectivity/no politics; documentation quality; workflow support	How participants described the operational value or limitations of the AI-PMS.
Emotional responses	Cold/lacks empathy; anxiety; discomfort; hope; motivation concerns	How participants emotionally reacted to AI-supported evaluation.
Task-related judgments	Trust in quantitative/KPI tasks; low trust in qualitative tasks; skepticism toward accuracy; task appropriateness	How participants differentiated between types of performance management tasks.
Transparency and understanding	Transparency need; unclear score meaning; explanation need; uncertainty about system logic	How participants understood, questioned, or struggled to interpret AI-generated outputs.
Human oversight expectations	Need for human oversight; manager review; hybrid/co-pilot language; override expectations	How participants described the desired role of human managers in AI-supported evaluation.

Fairness and bias perceptions	Fairness doubts; hidden bias; programmer/data bias; objectivity claims; reward inflexibility	How participants evaluated fairness, bias, and consistency in AI-supported decisions.
Developmental and future-use perceptions	Development guidance lacking; personalized growth needs; adaptability/future improvement; usability comments	How participants assessed the system's developmental value and future improvement potential.

Phase 3 – Searching for Themes

The development of themes involved an iterative and interpretive comparison across semantic and latent codes to identify broader patterns of meaning regarding employees' perceptions of the AI-based performance management system. Rather than grouping codes solely according to frequency or surface similarity, the analysis focused on how participants constructed tensions, contradictions, and evaluative judgments regarding AI-supported performance management processes.

An important interpretive pattern emerged around the simultaneous appreciation of operational efficiency and discomfort with the emotional limitations of AI-supported evaluation. Codes related to efficiency, speed, consistency, and automation repeatedly co-occurred with descriptions of emotional distance, impersonal feedback, and surveillance-like monitoring. These interconnected meanings contributed to the initial development of the theme “Efficient but Emotionally Cold,” which was later refined into “Emotionally-Disconnected Efficiency.”

One participant described this tension directly:

“The system is definitely faster and more organized than traditional evaluation methods. Everything is documented, and managers cannot easily ignore performance data. But at the same time, it feels emotionally distant. Sometimes it feels like the system sees employees as numbers rather than people.” (I9)

Another participant similarly emphasized the contrast between operational consistency and relational understanding:

“AI can monitor performance continuously and objectively, but it cannot understand the human side of work. Stress, personal difficulties, motivation, or emotional pressure cannot always be translated into data.” (I4)

A second interpretive cluster emerged around differentiated trust toward AI-supported evaluation. Participants frequently expressed higher confidence in AI when the system handled quantitative, structured, and measurable tasks, such as KPI tracking, reporting, or data analysis. However, trust weakened considerably when AI-supported systems were perceived as evaluating qualitative, relational, or developmental dimensions of performance. These tensions contributed to the development of the provisional theme “Conditional Trust,” which was later refined into “Task-Bound Trust.”

As one participant explained:

“For numerical evaluations or tracking performance indicators, AI is probably more reliable than humans because it processes data consistently. But when performance evaluation requires interpretation, mentoring, or understanding people, I still trust human judgment more.” (I15)

Similarly, concerns regarding transparency, hidden bias, and the need for managerial interpretation converged around participants’ expectations for meaningful human involvement in AI-supported evaluation processes. Participants generally preferred hybrid arrangements in which AI provided analytical support while managers retained responsibility for contextual interpretation and final evaluation decisions. These patterns contributed to the development of the theme “Human-in-the-Loop Safeguard.”

One participant explained:

“AI can support managers, but it should not become the final decision-maker. Employees need to know that a human being still reviews the process and understands the broader context behind the evaluation.” (I17)

At the same time, participants described complex and sometimes contradictory perceptions regarding fairness. Some employees perceived AI-supported evaluation as more objective and less vulnerable to favoritism compared to traditional manager-centered evaluation systems. Others expressed concern that hidden biases could still exist within algorithms, training data, or system design. These tensions contributed to the development of the provisional theme “Fairness Paradox,” which was later refined into “Algorithmic Fairness Tension.”

As one participant explained:

“The system may reduce personal favoritism, but algorithms are still designed by people. If the data or assumptions behind the system are biased, the results may also become unfair in a different way.” (I11)

Table A2 summarizes the preliminary thematic structure developed during this analytical stage. The table presents provisional themes, illustrative code groups, and the interpretive tensions that informed later theme refinement. These provisional themes were not treated as final themes at this stage; they were reviewed and refined in subsequent phases through comparison with coded extracts, analytic memos, and the research question.

Table A2. Preliminary Thematic Structure

Provisional theme	Illustrative code groups	Interpretive tension informing refinement
A1. Efficient but emotionally cold	Efficiency/time-saving; objectivity; documentation quality; cold/lacks empathy; context blindness	Participants valued AI-PMS for speed, consistency, and documentation, but also described

		it as emotionally distant and limited in understanding human circumstances.
A2. Conditional trust	Trust in quantitative/KPI tasks; lower trust in qualitative or relational tasks; skepticism toward interpretive accuracy	Participants trusted AI more for structured and measurable evaluation tasks, but less for developmental, relational, or context-sensitive judgments.
A3. Demand for human oversight	Need for human review; transparency need; unclear score meaning; manager interpretation; override expectations	Participants expected AI to support analysis, but emphasized that human managers should contextualize outputs and retain final evaluative authority.
A4. Fairness paradox	Objectivity/no favoritism; fairness doubts; hidden bias; programmer/data bias; equal versus equitable treatment	Participants perceived AI as potentially reducing personal favoritism while also raising concerns that hidden assumptions or biased data could create new forms of unfairness.

Phase 4 – Reviewing Themes

Themes were cross-checked against every coded extract and the full data set to ensure internal coherence and external distinctiveness.

Level-1 Check (coherence within themes)

- A1 was internally coherent: every extract contrasted speed/objectivity with lack of warmth.
- A2 cleanly separated high trust (quant tasks) from low trust (qual tasks).
- A3 mixed process-transparency and human-oversight; these were merged because both revolve around “keeping AI in check.”
- A4 contained two interconnected perceptions: AI was seen as potentially more objective than human evaluation, yet also vulnerable to hidden bias embedded in data, assumptions, or system design. Rather than treating these as separate themes, they were integrated into the final theme “Algorithmic Fairness Tension.”

Level-2 Check (distinctiveness across full data)

A2 and A4 overlapped in relation to bias risk; however, A2 concerned trust judgments, whereas A4 concerned perceptions of fairness outcomes. Both themes were therefore retained.

Resulting refined theme structure:

1. Emotionally-Disconnected Efficiency
2. Task-Bound Trust
3. Human-in-the-Loop Safeguard

4. Algorithmic Fairness Tension

Phase 5 – Defining and Naming Themes

Each theme was given a concise analytical definition and linked to the narrative evidence presented in Phase 3 (Table A3).

Table A3. Final Theme Definitions and Narrative Evidence

Final Theme	Formal Definition	Narrative Evidence
Theme 1: Emotionally-Disconnected Efficiency	The perception that AI-PMS improves speed, consistency, documentation, and objectivity while weakening the relational, emotional, and motivational qualities expected in performance feedback.	Supported by the narrative quotations from I9 and I4.
Theme 2: Task-Bound Trust	The perception that trust in AI-PMS depends on the type of evaluative task, with higher trust in quantitative, KPI-based, and structured tasks and lower trust in qualitative, relational, and developmental judgments.	Supported by the narrative quotation from I15.
Theme 3: Human-in-the-Loop Safeguard	The expectation that human managers should review, contextualize, and, when necessary, override AI-generated outputs to ensure interpretive accuracy, empathy, and accountability.	Supported by the narrative quotation from I7.
Theme 4: Algorithmic Fairness Tension	The simultaneous perception that AI-PMS may reduce overt favoritism while also creating concerns about hidden bias, opaque assumptions, and unfair outcomes embedded in data or system design.	Supported by the narrative quotation from I11.

Phase 6 – Producing the Report

In the final reporting phase, the four themes were organized into an integrated analytical narrative explaining how employees perceived AI-supported performance management. To maintain a consistent evidentiary chain from data to theme, the final report relied on the same verbatim quotations presented in Phase 3. Emotionally-disconnected efficiency was grounded in participants' accounts of AI as operationally efficient but relationally limited (I9, I4). Task-bound trust was grounded in differentiated trust between quantitative and qualitative evaluative tasks (I15). Human-in-the-loop safeguard was grounded in participants' expectation that managers should review, contextualize, and retain final interpretive authority over AI-generated outputs (I7). Algorithmic fairness tension was grounded in the simultaneous perception that AI may reduce favoritism while also reproducing hidden bias through data, assumptions, or system design (I11).

Finally, Table A4 reports illustrative high-frequency initial codes for analytical transparency. These descriptive frequencies were not used to determine theme importance.

Table A4. Illustrative High-Frequency Initial Codes

Code (Level-1)	Description	n
----------------	-------------	---

Objectivity / No Politics	Any statement that AI is impartial, consistent or free of personal bias.	38
Efficiency / Time-Saving	Mentions of speed, workload reduction, automation, or “makes work easier.”	34
Transparency Need	Requests to know how the algorithm works or how decisions are made.	29
Cold / Lacks Empathy	Descriptions of AI feedback as harsh, accusatory, emotionless, “surveillance.”	27
Need Human Oversight	Calls for human review, hybrid, co-pilot, or complement language.	24
Task-Bound Trust – Quant	High trust when AI handles quantitative or KPI tasks.	22
Task-Bound Trust – Qual	Low trust when AI handles qualitative, relational tasks.	20
Fairness Doubts	Concerns about hidden bias, programmer prejudice, or data-bias.	18
Development Guidance Lacking	AI fails to give personalized growth paths or mentoring.	16
Adaptability / Future Improvement	Hope that AI will improve with updates and learning.	15
Psychological Capital – Hope	Explicit hopefulness about AI impact.	14
Psychological Capital – Anxiety	Anxiety, discomfort, or fear about AI monitoring or job security.	11
User-Friendliness	Positive comments on UI, visuals, ease of use.	10
Context Blindness	AI ignores qualitative context, unique cases, or human factors.	10

About Alireza Ghezel (Corresponding Author):

Alireza Ghezel is a human resource management specialist and organizational behavior researcher with extensive practical and academic experience in performance management, talent development, leadership development, and AI-supported HR systems. He is a doctoral researcher in Management and Organizational Behavior, with a strong applied background in designing and implementing evidence-based performance management systems, competency models, assessment and development centers, and digital HR initiatives. His professional work focuses on bridging academic research and organizational practice, particularly in the areas of AI-based performance management, employee perceptions of algorithmic systems, human-centered HR technologies, and leadership development. He has contributed to multiple organizational consulting projects and academic studies related to sustainable HRM, digital transformation, coaching, and technology-enabled people management.

