IJ-ER
INDONESIAN JOURNAL ECONOMIC REVIEW

Received: 25 Jan 2026 | Revised: 31 Jan 2026 | Accepted: 15 Feb 2026 | Published: 28 Mar 2026 E-ISSN: 2986-478X, P-ISSN: 2986-1535

DOI: 10.59431/ijer.v6i1.729.

RESEARCH ARTICLE

Transparency, Algorithmic Fairness, and Employee Performance in AI-Based HR: The Moderating Role of Trust in Management

Muhammad Irvan [1*] | Bachrudin Syahroni [2] | Moh. Yudi Mahadianto [3]

[1*,2] Management Study Program, The Faculty of Economics and Business, Universitas Swadaya Gunung Jati, Cirebon City, West Java Province, Indonesia., Indonesia.

Correspondence
[1*] Management Study Program, The Faculty of Economics and Business, Universitas Swadaya Gunung Jati, Cirebon City, West Java Province, Indonesia., Indonesia.
Email. mohyudim@ugj.ac.id

Funding information
Universitas Swadaya Gunung Jati.

Abstract

This study examines how perceived transparency and algorithmic fairness in AI-based human resource (HR) practices relate to employee performance, and whether trust in management moderates these relationships. Data were collected through a quantitative cross-sectional survey of 152 employees from Indonesian service organizations using AI-enabled tools in HR decision processes, and analyzed using partial least squares structural equation modeling (PLS-SEM) with bootstrapping. The results show that perceived algorithmic fairness and trust in management have positive and significant relationships with employee performance, whereas perceived transparency does not have a significant direct effect. Trust in management also strengthens the positive relationship between perceived algorithmic fairness and employee performance, but does not significantly moderate the relationship between perceived transparency and employee performance. These findings suggest that organizations seeking performance benefits from AI-based HR should prioritize fairness governance and trust-building managerial practices, while designing transparency as meaningful, employee-relevant explanations rather than merely technical disclosure.

Keywords
Algorithmic Fairness; AI-Based HRM; Transparency; Trust In Management; Employee Performance; PLS-SEM.

1 | INTRODUCTION

Artificial intelligence (AI) is increasingly transforming how organizations manage human capital, from talent sourcing and screening to performance monitoring and people analytics. In principle, AI-enabled human resource management (HRM) can improve the speed, consistency, and scalability of HR decisions while supporting data-driven productivity gains (Tambe et al., 2019). However, many AI-based HR practices operate as "black-box" systems, making it difficult for employees and managers to understand what data are used, how decisions are generated, and why particular HR outcomes occur (Langer & König, 2023; Meijerink et al., 2021). This opacity is especially important in HR settings because such decisions affect employees' opportunities, rewards, and work experiences, all of which may shape individual performance and broader organizational outcomes.

A central challenge in AI-based Human Resources (HR) lies in the tension between the promise of objectivity and the potential for bias and perceived unfairness. While organizations may adopt AI to reduce human subjectivity in decision-making, there is a risk that algorithmic systems may inadvertently reproduce or even amplify existing biases. This can occur through biased training data, problematic feature selection, and unequal error rates across different groups (Mehrabi et al., 2022). In the context of HR, these risks are especially consequential, as employees not only assess the outcomes of decisions but also judge the fairness of the processes and the transparency of the explanations behind them. Research has shown that the use of algorithmic decision-making in HR can raise significant concerns about fairness, particularly when transparency and accountability mechanisms are insufficient (Kochling & Wehner, 2020; Newman et al., 2020). These concerns are not just ethical issues; they can have a direct impact on employees' acceptance of AI-driven HR practices. When employees perceive AI-driven decisions as unfair or opaque, it can negatively affect their willingness to cooperate with these systems and reduce their motivation to perform effectively. As AI becomes increasingly integrated into HR practices, it is critical to address these concerns to ensure the systems are perceived as just and reliable by all employees.

Drawing on Organizational Justice Theory, this study examines two governance-related perceptions that are particularly significant in AI-based Human Resources (HR) environments: perceived transparency and perceived algorithmic fairness. Transparency refers to the extent to which employees believe that AI-supported HR processes are understandable and that the criteria for decisions, data usage, and the logic behind those decisions are communicated in a clear and accessible manner (Langer & König, 2023). In contrast, algorithmic fairness pertains to employees' perceptions that AI-driven HR decisions are unbiased, consistent, and in line with widely accepted justice norms. Previous research has shown that perceptions of fairness remain a crucial psychological mechanism through which algorithmic management impacts employee attitudes and work-related outcomes. This is particularly true in HR contexts, where employees' trust and satisfaction can be influenced by how fair and transparent AI systems are perceived to be. Furthermore, studies suggest that transparency can only be effective if it provides meaningful explanations of decision-making processes, rather than simply disclosing information without clarity or relevance (Jabagi et al., 2025; Mirbabaie et al., 2025; Yu & Li, 2022). As AI systems continue to play a larger role in HR practices, understanding the importance of transparency and fairness becomes essential for fostering positive employee attitudes and ensuring the success of AI-driven HR initiatives.

In addition to transparency and fairness, employees' trust in management plays a crucial role in shaping how AI-based HR practices are interpreted and responded to. Trust theory suggests that employees are more willing to accept vulnerability when they have positive expectations about management's competence, integrity, and intentions (Mayer et al., 1995). In AI-driven HR environments, employees often do not see technological systems as neutral tools; instead, they view them as extensions of managerial priorities and organizational control (Jarrahi et al., 2021; Meijerink et al., 2021). When trust in management is high, employees are more likely to see AI-based HR practices as legitimate, supportive, and beneficial to their professional growth. However, when trust is low, employees may view these same practices as forms of surveillance or manipulation. Even with transparent and fair systems in place, a lack of trust can lead to negative perceptions. Trust, therefore, influences not only performance but also determines whether employees accept AI-driven decisions. Building and maintaining trust is essential for organizations looking to successfully implement AI-based HR systems and keep employees engaged.

Despite the growing global interest in algorithmic Human Resource Management (HRM), there remains a lack of micro-level empirical evidence on how perceived transparency and algorithmic fairness impact employee performance, particularly in emerging economies. In these contexts, digital HR transformation is advancing, but governance maturity is still uneven. In Indonesia, previous studies on HR digitalization, such as HR Information Systems (HRIS) and electronic HRM (e-HRM), suggest that technology-enabled HR systems are positively related to performance outcomes (Burhanuddin et al., 2026; Savitri et al., 2024). However, this body of research generally fails to address the unique challenges associated with AI-based HR systems, including issues such as opacity, algorithmic bias, and fairness perceptions. Furthermore, it does not provide clarity on when these perceptions are more likely to lead to improved employee performance. A key gap in the literature is the limited attention given to the role of trust in management, specifically whether it strengthens or weakens the effects of transparency and fairness in AI-based HR environments.

DOI: https://doi.org/10.59431/ijer.v6i1.729.

Given the growing use of AI in HR practices, it is essential to understand how these factors interact and influence employee outcomes in both mature and emerging economy settings.

To address these gaps, this study develops and tests a model linking perceived transparency of AI-based HR practices and perceived algorithmic fairness of HR decisions to employee performance, while also examining trust in management as both a direct predictor and a moderator. Using survey data from employees in Indonesian service organizations that use AI-enabled HR tools and analyzing the model with PLS-SEM, this study contributes theoretically by integrating algorithmic HR governance constructs with trust to explain employee performance. Practically, it offers guidance for organizations seeking to deploy AI-based HR in ways that improve performance while sustaining perceived fairness, legitimacy, and managerial credibility.

2 | BACKGROUND THEORY

Algorithmic HRM and AI-Based HR Practices

The use of artificial intelligence (AI) in human resource management (HRM) is increasingly described as algorithm-based or AI-assisted HRM, in which computational systems support or automate activities such as screening, selection, evaluation, and performance-related feedback. These systems are often valued because they can improve speed, consistency, scalability, and data-processing capacity compared with purely human decision-making (Tambe et al., 2019). However, AI-based HR practices also reshape organizational control, information asymmetry, and accountability, particularly when employees have limited visibility into how HR decisions are produced (Jarrahi et al., 2021; Langer & König, 2023). Accordingly, the organizational effects of AI-based HR cannot be understood solely in technical terms; they also depend on how employees interpret the legitimacy, fairness, and credibility of algorithm-supported HR decisions.

This issue is especially important in emerging-economy settings, where digital HR transformation may advance faster than the organizational routines, governance practices, and employee-facing safeguards needed to make AI-supported decisions broadly understandable and trusted. In such contexts, employees' perceptions of transparency, fairness, and managerial intent may become especially consequential for whether AI-enabled HR practices are accepted and translated into positive work outcomes.

Organizational Justice as the Core Explanatory Lens

Organizational Justice Theory provides the main explanatory lens for understanding employee reactions to AI-based HR practices. Justice theory emphasizes that employees evaluate not only decision outcomes, but also the fairness of the procedures used, the adequacy of explanations provided, and the degree to which decision processes are unbiased and respectful (Colquitt, 2001). A foundational meta-analysis demonstrates that justice perceptions are systematically associated with key workplace attitudes and behaviors, including performance-related outcomes (Cohen-Charash & Spector, 2001). This makes justice theory particularly relevant for AI-based HR, because algorithmic systems directly affect both the decision procedure and the informational environment surrounding HR decisions.

In AI-enabled HR contexts, employees may judge not only whether a decision benefits or disadvantages them, but also whether the decision process is understandable, consistent, and based on appropriate information. Prior studies show that algorithmic HR decisions can be perceived as less procedurally fair than equivalent human decisions when employees feel that algorithms reduce complex contributions into narrow or decontextualized indicators (Newman et al., 2020). Research on intelligent systems in organizational decision-making similarly shows that procedural justice remains essential regardless of whether decisions are made by humans or machine-supported systems (Ötting & Maier, 2018). Evidence from Indonesian organizational settings also supports the importance of justice-based mechanisms for employee outcomes, even outside explicitly AI-enabled HR environments (Mahadianto et al., 2025). Thus, justice theory offers a strong conceptual basis for explaining why governance-related perceptions in AI-based HR may influence employee performance.

Transparency in AI-Based HR Practices

Transparency is commonly understood as a governance mechanism for reducing opacity and strengthening accountability in AI-supported decision systems. In HR settings, however, transparency should not be interpreted merely as technical disclosure about models or data. Rather, it concerns whether employees receive explanations that are meaningful, timely, and appropriate to their role in the decision process (Langer & König, 2023). A transparency-by-design perspective further suggests that transparency should be embedded into automated decision systems from the outset in order to support legitimacy and responsible use, rather than added only after problems emerge (Felzmann et al., 2020).

From a justice perspective, perceived transparency may enhance employee responses when it reduces uncertainty, clarifies decision criteria, and helps employees understand how AI-supported HR decisions are generated. In principle, such clarity can strengthen perceptions of procedural and informational justice, increase acceptance of HR processes, and support more constructive work behavior. Yet transparency does not automatically produce positive outcomes. Recent research in algorithmic management points to a possible "transparency fallacy," whereby disclosure does not necessarily improve fairness perceptions and may even redirect attention to new concerns (Mirbabaie et al., 2025). Related evidence also suggests that AI decision transparency can simultaneously increase trust through perceived effectiveness while decreasing trust through discomfort or threat perceptions (Yu & Li, 2022). Likewise, AI-generated feedback may improve performance in use, but explicit disclosure that feedback is AI-generated can trigger adverse reactions (Tong et al., 2021). These findings imply that transparency is most likely to matter when employees perceive it as a credible and useful explanation rather than as disclosure for its own sake. Based on this reasoning, perceived transparency is expected to have a positive relationship with employee performance.

Algorithmic Fairness as a Governance-Relevant Perception

Algorithmic fairness refers to the extent to which AI-supported HR decisions are perceived as unbiased, consistent, and aligned with accepted fairness norms. At the technical level, unfairness may arise from biased data, problematic feature selection, model design choices, or deployment conditions that produce systematically unequal outcomes across groups (Mehrabi et al., 2021). In HR contexts, systematic review evidence indicates that algorithmic decision-making can create discrimination risks and fairness concerns in recruitment and development, while practical organizational responses to such issues remain underdeveloped relative to the technical fairness literature (Kochling & Wehner, 2020). As a result, even highly efficient systems may be viewed negatively if employees perceive them as reductionistic, unaccountable, or unjust (Newman et al., 2020). In organizations, fairness is experienced primarily as perceived fairness, that is, employees' subjective evaluation of whether decisions and procedures conform to accepted standards of justice. In AI-based HR, such evaluations may depend on the perceived consistency of evaluation rules, the relevance and completeness of information used, the absence of bias, and the availability of meaningful recourse or explanation (Colquitt, 2001; Newman et al., 2020). Empirical evidence from algorithm-mediated work settings shows that perceived fairness of algorithmic HR decisions is associated with important employee outcomes such as job satisfaction and perceived organizational support (Jabagi et al., 2025). Although some of this evidence comes from platform or digitally mediated work environments, its theoretical implication extends to broader organizational settings: fairness is a central psychological mechanism through which algorithmic management shapes cooperation, reciprocity, and productive work behavior. This argument is especially salient in emerging-economy contexts, where formal governance routines around AI may still be evolving and employees may rely more heavily on fairness signals to judge whether technologically mediated HR decisions are legitimate. When employees perceive AI-supported HR decisions as fair, they are more likely to see the organization as acting responsibly and to respond with stronger engagement and performance. Therefore, perceived algorithmic fairness is expected to have a positive relationship with employee performance.

Trust in Management as a Complementary Mechanism and Boundary Condition

While transparency and fairness characterize the decision environment, employees often interpret AI-based HR practices as reflections of managerial intent and organizational priorities. Trust theory defines trust as a willingness to accept vulnerability based on positive expectations regarding another party's competence, benevolence, and integrity (Mayer et al., 1995). In organizational settings, trust in leadership has been consistently linked to attitudes and behaviors relevant to performance, including cooperation, commitment, and extra effort (Dirks & Ferrin, 2002). This suggests that trust in management may itself directly support higher employee performance.

Trust is also likely to shape how employees interpret governance signals in AI-based HR. Because AI systems are deployed, configured, and authorized by management, employees may not see them as neutral tools, but as extensions of managerial control and intent (Jarrahi et al., 2021; Meijerink et al., 2021). Under high trust, employees may interpret transparency and fairness as credible signals that the organization is competent, responsible, and supportive. Under low trust, the same practices may be interpreted more skeptically, for example as surveillance, manipulation, or symbolic compliance. In this sense, trust in management functions not only as a direct antecedent of employee performance but also as a boundary condition that may strengthen the positive effects of transparency and fairness on performance.

Employee Performance in the Context of AI-Based HR

Employee performance is commonly conceptualized as a multidimensional construct that includes task performance, contextual performance, and the avoidance of counterproductive behavior. The Individual Work Performance Questionnaire (IWPQ) is one of the most widely used approaches for capturing these dimensions across diverse job contexts (Koopmans et al., 2013). Evidence for the Indonesian adaptation of the IWPQ also supports its relevance for local organizational research (Widyastuti et al., 2024). This is important because AI-based HR practices may influence not only how employees complete formal job tasks, but also how they engage in discretionary behaviors such as initiative, cooperation, and helping. From the standpoint of digital HR transformation in Indonesia, prior studies on HRIS and e-HRM suggest that technology-enabled HR systems can relate positively to employee performance and productivity (Burhanuddin et al., 2026; Nirwana et al., 2023; Sumaryono, 2023). However, this stream of research generally does not address the distinctive governance challenges of AI-based HR, including opacity, algorithmic bias, and fairness perceptions. For that reason, employee performance is an appropriate and meaningful outcome for examining whether transparency, algorithmic fairness, and trust in management jointly shape how employees respond to AI-supported HR environments.

Hypotheses Development

Building on Organizational Justice Theory (Colquitt, 2001) and social exchange theory, which links fair treatment and trustworthy leadership to reciprocal work behaviors (Cohen-Charash & Spector, 2001; Cropanzano & Mitchell, 2005; Dirks & Ferrin, 2002), this study proposes the following hypotheses: H1 suggests that the perceived transparency of AI-based HR practices positively influences employee performance. H2 posits that the perceived algorithmic fairness of HR decisions positively influences employee performance. H3 hypothesizes that trust in management positively influences employee performance. H4 examines whether trust in management moderates the relationship between perceived transparency and employee performance, with the relationship being stronger when trust is higher. Finally, H5 proposes that trust in management also moderates the relationship between perceived algorithmic fairness and employee performance, such that the relationship is stronger when trust is higher.

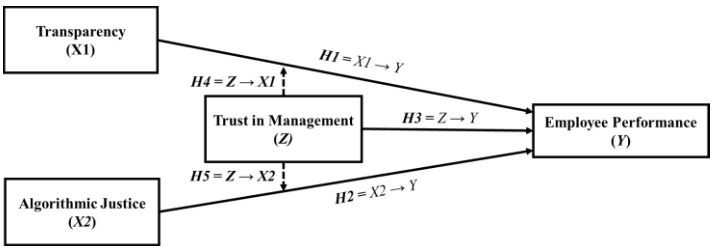

Figure 1. Research Model

3 | METHOD

This study employed a quantitative, explanatory research design utilizing a cross-sectional online survey to test the proposed relationships among perceived transparency of AI-based HR practices, perceived algorithmic fairness, trust in management, and employee performance. A survey design was appropriate as the focal constructs represent employees' perceptions and self-reported work behaviors, and the research objective was to estimate predictive relationships, interactions, and moderation effects within a latent-variable model. To analyze the predictive model with moderation, Partial Least Squares Structural Equation Modeling (PLS-SEM) was used, which is suitable for prediction-oriented models with latent constructs and interaction terms (Hair Jr. et al., 2019; Henseler et al., 2016).

The target population consisted of employees working in Indonesian service-sector organizations that use AI-enabled tools in HR-related decision processes, such as screening, employee evaluations, performance-related recommendations, or other decision-support functions. Since a comprehensive sampling frame for such employees was not available, a non-probability purposive sampling approach was adopted (Etikan et al., 2016). This approach was deemed appropriate as the study required respondents who met specific inclusion criteria relevant to the research context. The eligibility criteria were: (1) currently employed in a service organization, (2) awareness or experience of AI systems being used in at least one HR-related decision or evaluation process within the organization, and (3) willingness to participate voluntarily. Data were collected via an online questionnaire distributed through organizational and professional networks. After data screening, the final dataset included 152 valid responses, which is generally sufficient for PLS-SEM applications with moderation when measurement quality is ensured and statistical power is considered (Hair Jr. et al., 2019). For planning purposes, statistical power was also assessed using a priori reasoning based on G*Power guidelines for regression-type models (Faul et al., 2009).

All constructs were operationalized as reflective multi-item measures and assessed using a 5-point Likert-type scale ranging from 1 = strongly disagree to 5 = strongly agree. The questionnaire was developed by adapting and contextualizing established measures to fit AI-enabled HR decision environments while maintaining conceptual coverage for each construct. Perceived Transparency of AI-based HR Practices (X1). Transparency was measured using 8 items that reflect employees' perceptions of AI-supported HR processes being explainable, of decision logic and relevant information being communicated, and of employees having clarity and avenues for raising concerns. Items were adapted from empirical operationalizations of transparency and fairness perceptions in algorithmic recruiting and related decision settings (Ochmann et al., 2024). Perceived Algorithmic Fairness (X2). Algorithmic fairness was measured using 8 items reflecting perceived consistency, lack of bias, the relevance of the data used, and the alignment between decisions and employees' contributions. The measurement approach aligns with justice-based conceptualizations of algorithmic fairness perceptions and is adapted from prior empirical work linking transparency interventions and fairness perceptions in algorithmic decision contexts (Ochmann et al., 2024). Trust in Management (Z). Trust in management was measured using 6 items representing employees' confidence in management's competence and integrity, as well as the belief that management acts with concern for employees. Items were adapted from a validated trust-for-management measurement used in organizational field research (Mayer & Davis, 1999). Employee Performance (Y). Employee performance was measured using 6 items that capture task and contextual performance aspects commonly used in individual work performance research. Items were adapted from individual work performance indicators and validated performance measurement approaches (Koopmans et al., 2014). To ensure local context relevance, the measurement approach was consistent with psychometric evidence for Indonesian versions of individual work performance measures (Widyastuti et al., 2024).

Since the study was conducted in Indonesian workplaces, the instrument was prepared in Bahasa Indonesia. A structured cross-cultural adaptation procedure involving translation, reconciliation, and review for conceptual equivalence was applied following established guidelines for adapting self-report measures (Beaton et al., 2000). Before full deployment, the questionnaire was also reviewed for face and content validity by individuals familiar with HR and organizational research to ensure item clarity and contextual appropriateness. The survey was administered online during a brief data collection period. Participation was voluntary, and respondents completed an informed consent statement before accessing the questionnaire. To improve response quality, the instructions emphasized honest reporting, clarified that there were no right or wrong answers, and informed respondents that they could withdraw at any time. Data screening procedures included checks for incomplete submissions and patterns indicating inattentive responding. Only responses meeting the inclusion criteria and completeness requirements were retained for analysis.

PLS-SEM analysis followed current reporting recommendations (Hair Jr. et al., 2019; Henseler et al., 2016) and proceeded in two stages. First, the measurement model was evaluated in terms of indicator reliability, convergent validity, internal consistency reliability, and discriminant validity. Indicator loadings and average variance extracted (AVE) were used to assess convergent validity, while Cronbach's alpha and composite reliability were used to assess internal consistency reliability (Hair Jr. et al., 2019; Tavakol & Dennick, 2011). Discriminant validity was assessed using the Fornell–Larcker criterion (Fornell & Larcker, 1981) and the heterotrait–monotrait ratio (HTMT) criterion (Henseler et al., 2015). Second, the structural model was assessed by examining collinearity, path coefficients, effect sizes, explanatory power, and model fit. Collinearity among predictors was evaluated using variance inflation factors (VIF). Path coefficients and interaction effects were estimated and tested for significance using bootstrapping within the PLS-SEM framework (Hair Jr. et al., 2019). Moderation hypotheses were tested by creating interaction terms between trust in management and each focal predictor, namely perceived transparency and perceived algorithmic fairness. To support fit information, the standardized root mean square residual (SRMR) was also examined using established guideline values (Hu & Bentler, 1999), while overall model quality was interpreted primarily based on measurement validity, reliability, and predictive structural assessment in line with PLS-SEM recommendations (Hair Jr. et al., 2019; Henseler et al., 2016). Since the study relied on self-report survey data, procedural remedies were implemented to reduce the risk of common method bias, including ensuring respondent anonymity, reducing evaluation apprehension, and emphasizing that there were no right or wrong answers (Podsakoff et al., 2003). As a complementary statistical diagnostic, full collinearity VIF values were also assessed as an indicator of potential common method bias contamination in PLS-SEM (Kock, 2015).

The study adhered to standard ethical principles for behavioral research. Respondents were informed about the purpose of the study, the voluntary nature of participation, confidentiality, and the intended use of the data before giving consent. No identifying personal information was collected, and all results were reported in aggregate form. Methodologically, this study is subject to several limitations. The cross-sectional design limits causal inferences, the use of self-reported measures may increase the potential for common method variance, and the non-probability purposive sampling strategy may limit generalizability. Nevertheless, these risks were addressed through procedural controls during data collection and complementary statistical diagnostics during analysis.

4 | RESULTS AND DISCUSSION

4.1 Results

4.1.1 Respondent characteristics

A total of 152 valid responses were included in the analysis. As shown in Table 1, the sample was relatively balanced by gender (53.95% male and 46.05% female), with the largest proportion of respondents aged 25–40 years (51.32%). Most respondents held a bachelor's degree (57.89%), and the largest tenure category was 4–6 years of work experience (46.71%). These characteristics indicate that the sample represents employees with sufficient educational and work exposure to evaluate AI-enabled HR practices in their organizations.

Table 1. Demographic characteristics of research participants

Variable	Category	n	%
Gender			
	Male	82	53.95
	Female	70	46.05
Age			
	< 25 years	46	30.26
	25–40 years	78	51.32
	41–50 years	28	18.42
Highest Education			
	Diploma	16	10.53
	Bachelor’s	88	57.89
	Master’s	48	31.58
Tenure			
	< 1 year	7	4.61
	1–3 years	64	42.11
	4–6 years	71	46.71
	> 6 years	10	6.58

Note. Percentages are based on N = 152 respondents. Categories with zero frequency are omitted for brevity.

Table 1 presents the demographic characteristics of the research participants. Among the 152 respondents, 53.95% were male and 46.05% female. In terms of age, 51.32% were between 25–40 years, while 30.26% were under 25 years. Regarding education, 57.89% had a bachelor’s degree. The majority had tenure between 1–6 years, with 46.71% having 4–6 years of experience. Categories with zero frequency were omitted for clarity.

4.1.2 Descriptive statistics

Table 2 presents the descriptive statistics for the study variables measured on a 5-point scale. Overall, respondents reported moderate levels of perceived transparency (M = 3.121, SD = 0.807), perceived algorithmic fairness (M = 3.214, SD = 0.834), trust in management (M = 3.201, SD = 0.817), and employee performance (M = 3.072, SD = 0.855). These values suggest that the focal perceptions and outcome variable were present at moderate levels in the sample, providing sufficient variation for subsequent structural analysis.

Table 2. Descriptive Statistics of Study Variables

Construct	N	Min	Max	Mean & SD
Transparency	152	1.5	5.0	3.121 (0.807)
Algorithmic fairness	152	1.125	5.0	3.214 (0.834)
Trust in management	152	1.0	5.0	3.201 (0.817)
Employee performance	152	1.167	5.0	3.072 (0.855)

Note. Scale anchors: 1 = strongly disagree, 5 = strongly agree.

Table 2 presents the descriptive statistics for the study variables. For 152 respondents, the mean scores for transparency, algorithmic fairness, trust in management, and employee performance ranged from 3.072 to 3.214, with standard deviations ranging from 0.807 to 0.855. The minimum scores varied from 1.0 to 1.5, while the maximum scores reached 5.0 for all constructs. These results reflect moderate perceptions across the constructs, with transparency showing the lowest mean and employee performance the lowest standard deviation.

4.1.3 Measurement model assessment

The reflective measurement model was assessed in terms of indicator reliability, internal consistency reliability, convergent validity, and discriminant validity following established PLS-SEM guidelines. As reported in Table 3, all indicator loadings exceeded the recommended threshold of 0.70, ranging from 0.714 to 0.871 for transparency, 0.766 to 0.846 for algorithmic fairness, 0.801 to 0.854 for trust in management, and 0.782 to 0.831 for employee performance. These results support adequate indicator reliability.

Table 3. Indicator loadings

Construct	Indicator	Loading
Algorithmic fairness	AF01	0.825
Algorithmic fairness	AF02	0.825

Algorithmic fairness	AF03	0.846
Algorithmic fairness	AF04	0.797
Algorithmic fairness	AF05	0.809
Algorithmic fairness	AF06	0.766
Algorithmic fairness	AF07	0.819
Algorithmic fairness	AF08	0.781
Employee performance	EP01	0.798
Employee performance	EP02	0.782
Employee performance	EP03	0.824
Employee performance	EP04	0.790
Employee performance	EP05	0.818
Employee performance	EP06	0.831
Transparency	TR01	0.766
Transparency	TR02	0.871
Transparency	TR03	0.858
Transparency	TR04	0.799
Transparency	TR05	0.867
Transparency	TR06	0.714
Transparency	TR07	0.750
Transparency	TR08	0.738
Trust in management	TM01	0.826
Trust in management	TM02	0.854
Trust in management	TM03	0.837
Trust in management	TM04	0.801
Trust in management	TM05	0.835
Trust in management	TM06	0.849

Note. All loadings are from the PLS-SEM algorithm estimates (SmartPLS 4).

Table 4 further shows that all constructs demonstrated strong internal consistency and convergent validity. Cronbach's alpha values ranged from 0.893 to 0.927, composite reliability values ranged from 0.918 to 0.938, and all average variance extracted (AVE) values exceeded 0.50, ranging from 0.636 to 0.695. Thus, all constructs satisfied commonly accepted reliability and convergent validity criteria.

Table 4. Construct reliability and convergent validity

Construct	Cronbach's alpha	rho_A	Composite reliability	AVE
Algorithmic fairness	0.925	0.928	0.938	0.654
Employee performance	0.893	0.896	0.918	0.652
Transparency	0.927	0.978	0.933	0.636
Trust in management	0.912	0.917	0.932	0.695

Note. rho_A = Dijkstra–Henseler's rho_A; AVE = average variance extracted.

Discriminant validity was examined using the heterotrait–monotrait ratio (HTMT). As shown in Table 5, all HTMT values were below 0.85, indicating adequate discriminant validity among the focal constructs. Taken together, these results indicate that the measurement model was satisfactory and suitable for structural model evaluation.

Table 5. Discriminant validity (HTMT)

Construct A	Construct B	HTMT
Employee performance	Algorithmic fairness	0.530
Transparency	Algorithmic fairness	0.103
Transparency	Employee performance	0.164
Trust in management	Algorithmic fairness	0.379
Trust in management	Employee performance	0.535
Trust in management	Transparency	0.194

Note. All HTMT values are below 0.85, supporting discriminant validity.

Table 5 presents the HTMT values for discriminant validity between constructs. The values range from 0.103 to 0.535, with all HTMT values below the threshold of 0.85, indicating adequate discriminant validity. The highest HTMT value was observed between trust in management and employee performance (0.535), while the lowest value occurred between transparency and algorithmic fairness (0.103). These results suggest that the constructs are sufficiently distinct from one another, supporting the validity of the structural model.

4.1.4 Structural model assessment

Potential multicollinearity in the structural model was examined using variance inflation factors (VIF). All predictor VIF values for employee performance were well below commonly used cutoffs (e.g., VIF < 3.3), suggesting that collinearity was not a concern (Kock, 2015) (Table 6). Before testing the hypotheses, potential multicollinearity in the structural model was assessed using variance inflation factors (VIF). As presented in Table 6, all VIF values were well below commonly used cutoffs, ranging from 1.057 to 1.225 for the main predictors and from 1.147 to 1.149 for the interaction terms. This indicates that collinearity was not a concern in the model.

Table 6. Inner model collinearity statistics (VIF)

Predictor	VIF
Algorithmic fairness	1.148
Transparency	1.057
Trust in management	1.225
Trust x Transparency	1.147
Trust x Algorithmic fairness	1.149

Note. VIF values are reported for predictors of employee performance.

The model accounted for 39.0% of the variance in employee performance (R^2 = 0.390; adjusted R^2 = 0.369), suggesting a moderate level of explanatory power, which is considered typical for behavioral research. This indicates that the independent variables included in the model provide a reasonable understanding of employee performance, although other factors not captured in the model may also influence performance. The adjusted R^2 value reflects a correction for the number of predictors, confirming that the model's explanatory power remains robust even after accounting for model complexity. These findings suggest that the model offers valuable insights into performance determinants.

Table 7. Coefficient of determination (R^2)

Endogenous construct	R-square	R-square adjusted
Employee performance	0.390	0.369

Note. R^2 values are for the endogenous construct employee performance.

Effect sizes (f^2) show that algorithmic fairness and trust in management had medium-sized contributions to employee performance (f^2 = 0.191 and 0.178, respectively), whereas transparency had a negligible effect (f^2 = 0.014). The interaction term trust × algorithmic fairness exhibited a small effect (f^2 = 0.041) (Table 8).

Table 8. Effect sizes (f^2) for predictors of employee performance

Predictor	f^2
Algorithmic fairness	0.191
Transparency	0.014
Trust in management	0.178
Trust x Transparency	0.009
Trust x Algorithmic fairness	0.041

Note. Effect size interpretation follows common guidelines (small ≈ 0.02, medium ≈ 0.15, large ≈ 0.35).

The overall model fit was assessed using the standardized root mean square residual (SRMR), a common indicator of model fit quality. The estimated SRMR value for the model was 0.070, which is below the recommended threshold of 0.08 (Hu & Bentler, 1999), indicating that the model exhibits an acceptable fit. This suggests that the model adequately represents the relationships between the variables and is consistent with established guidelines for structural equation modeling. Therefore, the model can be considered reliable for testing the proposed hypotheses.

Table 9. Model fit indices

Index	Saturated model	Estimated model
SRMR	0.071	0.070
d_ULS	2.019	2.016
d_G	0.565	0.565

Chi-square	463.124	462.279
NFI	0.845	0.846

Note. SRMR = standardized root mean square residual; d_ULS = squared Euclidean distance; d_G = geodesic distance; NFI = normed fit index.

The structural model explained 39.0% of the variance in employee performance (R^2 = 0.390; adjusted R^2 = 0.369), indicating moderate explanatory power for behavioral research (Table 7). Effect size estimates reported in Table 8 further show that algorithmic fairness (f^2 = 0.191) and trust in management (f^2 = 0.178) made medium-sized contributions to employee performance, whereas transparency had a negligible effect (f^2 = 0.014). The interaction term trust × algorithmic fairness showed a small effect (f^2 = 0.041), while trust × transparency had a negligible effect (f^2 = 0.009). In addition, the model demonstrated acceptable fit, with an estimated SRMR value of 0.070, which is below the 0.08 guideline (Table 9).

4.1.5 Hypothesis testing and moderation effects

Hypotheses were tested using bootstrapping, and the results are summarized in Table 10. The findings show that perceived transparency did not have a significant direct effect on employee performance (β = 0.094, t = 0.981, p = 0.326), indicating that H1 was not supported. In contrast, perceived algorithmic fairness had a positive and significant effect on employee performance (β = 0.366, t = 6.134, p < 0.001), supporting H2. Trust in management also had a positive and significant direct effect on employee performance (β = 0.365, t = 5.315, p < 0.001), supporting H3.

Table 10. Hypothesis testing results (bootstrapping)

Hyp.	Path	Beta	t	p	CI	f^2	Decision
H1	Transparency on Employee performance	0.094	0.981	0.326	[-0.194, 0.245]	0.014	Not supported
H2	Algorithmic fairness on Employee performance	0.366	6.134	< .001	[0.245, 0.481]	0.191	Supported
H3	Trust in management on Employee performance	0.365	5.315	< .001	[0.222, 0.487]	0.178	Supported
H4	Trust × Transparency on Employee performance	-0.076	0.950	0.342	[-0.226, 0.088]	0.009	Not supported
H5	Trust × Algorithmic fairness on Employee performance	0.152	2.453	0.014	[0.025, 0.269]	0.041	Supported

Note. Standardized coefficients are reported. CI = 95% percentile bootstrap confidence interval.

With respect to the moderating effects, the interaction between trust in management and perceived transparency was not significant (β = −0.076, t = 0.950, p = 0.342), so H4 was not supported. However, the interaction between trust in management and perceived algorithmic fairness was positive and significant (β = 0.152, t = 2.453, p = 0.014), supporting H5. These results indicate that trust in management strengthens the positive relationship between perceived algorithmic fairness and employee performance, but does not significantly alter the relationship between perceived transparency and employee performance. Figure 2 presents the estimated structural model, including the significant direct paths and the significant interaction term. Employee performance exhibited R^2 = 0.390.

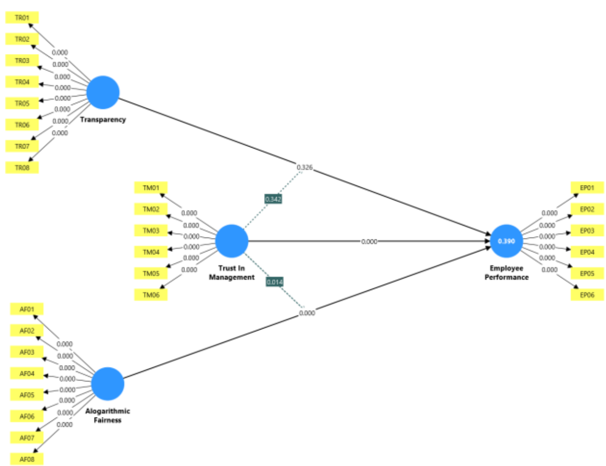

Figure 2. Structural model estimates from SmartPLS 4 (standardized path coefficients).

To better illustrate the significant moderation effect of hypothesis H5, Figure 3 depicts the conditional relationship between perceived algorithmic fairness and employee performance at varying levels of trust in management. Specifically, it shows the relationship at low and high trust levels (−1 SD and +1 SD), using standardized coefficients to highlight how trust in management influences the strength of this relationship. This visual representation helps to clarify the moderating role of trust in management on the impact of algorithmic fairness on employee performance, offering deeper insights into the dynamics between these variables.

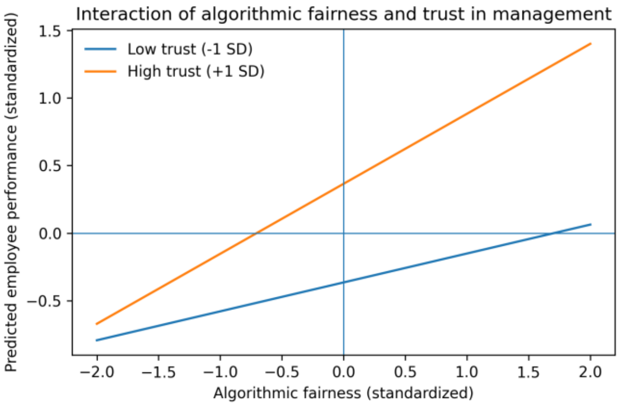

Figure 3. Simple slope plot for trust in management moderating the effect of algorithmic fairness on employee performance.

4.2 Discussion

The findings provide four substantive insights into how employees respond to AI-based HR practices. First, perceived algorithmic fairness showed a positive and statistically significant relationship with employee performance. This result reinforces Organizational Justice Theory, which suggests that employees are more likely to respond constructively when decision procedures are perceived as unbiased, consistent, and normatively appropriate (Colquitt, 2001). In AI-enabled HR settings, this mechanism is especially salient because employees may perceive algorithmic systems as reductionistic or decontextualized when they cannot see how contextual factors are incorporated into personnel decisions (Newman et al., 2020). The present finding therefore suggests that fairness perceptions remain a central behavioral pathway through which AI-supported HR governance influences performance. In practical terms, employees appear more willing to reciprocate and sustain productive work behavior when they believe AI-based HR decisions are applied fairly.

Second, trust in management was also positively related to employee performance. This finding is consistent with trust theory, which posits that trust facilitates cooperation, reduces uncertainty, and encourages employees to exert effort when organizational processes affect their interests (Mayer et al., 1995). It also aligns with broader evidence that trust in leadership is associated with performance-relevant attitudes and behaviors (Dirks & Ferrin, 2002). In the context of AI-based HR, this result is particularly important because employees are unlikely to interpret algorithmic systems as neutral or self-governing; rather, they tend to view them as extensions of managerial intent and organizational priorities. Accordingly, the performance consequences of AI-enabled HR practices depend not only on the system itself, but also on whether employees believe management is competent, fair, and acting in good faith when deploying such systems.

Third, trust in management positively moderated the relationship between perceived algorithmic fairness and employee performance. The simple slope pattern indicates that the positive effect of fairness becomes stronger when employees report higher trust in management. This finding extends the justice-based explanation by showing that fairness signals do not operate in isolation; they become more behaviorally consequential when employees view managerial intentions and competence as credible. In other words, trust functions as a boundary condition that amplifies the motivational value of perceived fairness. This is an important contribution because it shows that AI-based HR governance is not merely a matter of improving system design or procedural rules. Even when employees perceive AI-supported HR decisions as fair, the performance benefits of that fairness are likely to be stronger when management is trusted to implement and oversee the technology responsibly.

Fourth, perceived transparency did not have a significant direct relationship with employee performance, and its interaction with trust in management was also not significant. This result suggests that transparency alone may be insufficient to generate performance gains in AI-based HR settings. One plausible explanation is that transparency does not automatically function as a positive governance signal unless it also enhances perceptions of fairness, credibility, or managerial intent. Prior research has shown that transparency can create mixed reactions: while disclosure may increase perceived effectiveness, it can also generate discomfort, threat perceptions, or distrust when employees interpret the disclosed information as confusing, excessive, or merely symbolic (Yu & Li, 2022). Related evidence also indicates that disclosing the AI origin of otherwise useful feedback can trigger adverse reactions in workplace settings (Tong et al., 2021). Taken together, these dynamics help explain why transparency in the present study did not translate directly into higher employee performance. The finding implies that not all governance-relevant signals in AI-based HR are equally performance-relevant; fairness appears to be more proximal to employee behavior than transparency by itself.

Taken as a whole, the study makes three theoretical contributions. First, it supports the use of Organizational Justice Theory and trust theory as complementary lenses for explaining employee responses to AI-based HR practices. Second, it shows that perceived algorithmic fairness is a more decisive predictor of employee performance than perceived transparency, suggesting that employees care more about whether AI-supported HR decisions are substantively just than whether the system is merely explained. Third, it demonstrates that trust in management does not strengthen all governance signals uniformly; rather, its moderating role appears selective, strengthening the effect of fairness but not that of transparency. This selective pattern advances the literature by showing that trust is not a general amplifier of AI governance features, but a contextual condition that matters more when employees evaluate the legitimacy of decision outcomes and procedures.

The findings also carry several practical implications for organizations implementing AI-based HR. First, organizations should prioritize fairness governance over disclosure-oriented transparency alone. This includes auditing datasets and models for bias, ensuring that decision rules are applied consistently, validating that evaluation criteria are relevant to actual job contributions, and providing recourse mechanisms when employees wish to question or challenge AI-supported HR decisions. Second, organizations should recognize that managerial trust is not peripheral to digital HR transformation; leadership competence, integrity, and communication practices shape whether employees interpret AI-based HR as supportive or controlling. Trust-building efforts such as clear accountability, responsive leadership communication, and visible oversight of AI-based HR decisions are therefore likely to improve the conditions under which fairness perceptions translate into stronger performance. Third, transparency initiatives should be designed as meaningful employee-facing explanations rather than merely technical disclosure. In practice, explanations should help employees understand the relevance of the data used, the rationale of decision criteria, and the appropriate channels available when they need clarification or contestation.

Several limitations should be considered when interpreting these findings. The cross-sectional design limits strong causal inference, so the reported relationships should be interpreted primarily as theoretically grounded associations rather than definitive causal effects. The use of nonprobability purposive sampling in Indonesian service organizations also limits generalizability across sectors, occupations, and national contexts. In addition, all variables were measured through self-report, which may reflect perceived performance more strongly than objective performance and may introduce common method variance, even though procedural remedies and collinearity diagnostics were applied. Future research could address these limitations by using longitudinal or experimental designs, incorporating supervisor-rated or objective performance indicators, and testing the model across broader sectors and cultural settings. It would also be valuable to examine additional mechanisms such as perceived procedural justice, explanation quality, algorithm aversion, privacy concerns, or contestability to clarify why transparency alone may not consistently translate into improved employee performance in AI-based HR environments.

5 | CONCLUSIONS AND FUTURE WORK

This study shows that in AI-based HR contexts, employee performance is shaped more strongly by perceived algorithmic fairness and trust in management than by transparency alone. The findings indicate that perceived algorithmic fairness and trust in management have positive and significant relationships with employee performance, whereas perceived transparency does not have a significant direct effect. In addition, trust in management strengthens the positive relationship between perceived algorithmic fairness and employee performance, but does not significantly moderate the relationship between perceived transparency and employee performance.These findings contribute to the literature on AI-based HR governance in three ways. First, they support the use of Organizational Justice Theory and trust theory as complementary lenses for explaining employee responses to AI-enabled HR practices. Second, they show that fairness is a more decisive predictor of employee performance than transparency, suggesting that employees respond more strongly to whether AI-supported HR decisions are perceived as substantively just than to whether the system is merely disclosed or explained. Third, they show that trust in management is a selective boundary condition: it strengthens the performance relevance of fairness, but not transparency.

For practice, the results suggest that organizations seeking performance benefits from AI-based HR should prioritize fairness governance and trust-building managerial practices over disclosure-oriented transparency alone. This includes auditing AI-supported HR decisions for bias, ensuring that evaluation criteria are relevant and consistently applied, providing meaningful avenues for clarification or appeal, and strengthening leadership communication and accountability so that employees perceive AI-based HR systems as credible and responsibly governed. Transparency remains important, but it should be designed as meaningful employee-facing explanation rather than as purely technical disclosure. This study has several limitations. The cross-sectional design limits strong causal inference, the nonprobability purposive sampling approach restricts generalizability, and the use of self-reported measures may capture perceived rather than objective performance while also raising the possibility of common method bias. Future research could address these limitations by using longitudinal or experimental designs, incorporating supervisor-rated or objective performance indicators, and examining additional mechanisms or boundary conditions such as explanation quality, procedural justice, privacy concerns, algorithm aversion, and contestability. Overall, the study suggests that the success of AI-based HR depends not only on making systems more visible, but more importantly on ensuring that they are perceived as fair and are implemented by management that employees trust.

REFERENCES

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191.

Burhanuddin, B., Narinawati, U., Kurniawan, B., Riyanto, A., & Santy, R. D. (2026). Pengaruh Human Resource Information System (HRIS) Terhadap Kinerja Sumber Daya Manusia Dengan Efektivitas Kerja Sebagai Variabel Intervening. *Syntax Literate: Jurnal Ilmiah Indonesia, 11*(2), 1125–1137. https://doi.org/10.36418/syntax-literate.v11i2.63828

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86*(2), 278–321. https://doi.org/10.1006/obhd.2001.2958

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386–400. https://doi.org/10.1037/0021-9010.86.3.386

Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management, 31*(6), 874–900. https://doi.org/10.1177/0149206305279602

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology, 87*(4), 611–628. https://doi.org/10.1037/0021-9010.87.4.611

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Science and Engineering Ethics, 26*(6), 3333–3361. https://doi.org/10.1007/s11948-020-00276-4

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 382–388. https://doi.org/10.2307/3150980

Hair Jr., J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jabagi, N., Croteau, J., Audebrand, C.-H., & Marsan, J. (2025). Do algorithms play fair? Analysing the perceived fairness of HR-decisions made by algorithms and their impact on job satisfaction and perceived organisational support. *The International Journal of Human Resource Management, 36*(2), 235–274. https://doi.org/10.1080/09585192.2024.2441448

Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society, 8*, 1–14. https://doi.org/10.1177/20539517211020332

Kochling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research, 13*(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration, 11*(4), 1–10. https://doi.org/10.4018/IJeC.2015100101

Koopmans, L., Bernaards, C., Hildebrandt, V., Schaufeli, W. B., De Vet, H. C. W., & Van der Beek, A. J. (2014). Measuring individual work performance: Identifying and selecting indicators. *Work, 48*(2), 229–238. https://doi.org/10.3233/WOR-131659

Koopmans, L., Bernaards, C., Hildebrandt, V., van Buuren, S., van der Beek, A. J., & de Vet, H. C. W. (2013). Development of an individual work performance questionnaire. *International Journal of Productivity and Performance Management, 62*(1), 6–28. https://doi.org/10.1108/17410401311285273

Langer, M., & König, C. J. (2023). Introducing a multi-stakeholder perspective on algorithm-based human resource management: A review and agenda for future research. *Human Resource Management Review, 33*(1), 100881. https://doi.org/10.1016/j.hrmr.2021.100881

Mahadianto, M. Y., Muhammad Siddiq, D., Mamat, M., Abu Bakar, H., & Hartono, E. (2025). How procedural and informational justice moderated by organizational collectivism predict employee organizational commitment. *International Research Journal of Business Studies, 18*(2), 159–173. https://doi.org/10.21632/irjbs.18.2.159-173

Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology, 84*(1), 123–136. https://doi.org/10.1037/0021-9010.84.1.123

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 115. https://doi.org/10.1145/3457607

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 115. https://doi.org/10.1145/3457607

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: Synthesizing developments and cross-disciplinary insights on digital HRM. *The International Journal of Human Resource Management, 32*(12), 2545–2562. https://doi.org/10.1080/09585192.2021.1925326

Mirbabaie, M., Langer, M., Rieskamp, J., & Hofeditz, L. (2025). Transparency fallacy: Perceived fairness in algorithmic management. *Business & Information Systems Engineering.* https://doi.org/10.1007/s12599-025-00963-1

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes, 160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Nirwana, N. K. A., Makduani, R., Marhum, U., Rahmayanti, R., & Widaningsih, R. R. A. (2023). Dampak implementasi HRIS (Human Resource Information System) terhadap kinerja karyawan. *Jurnal Darma Agung, 31*(6), 393–400.

https://doi.org/10.46930/ojsuda.v31i6.3685

Ochmann, J., Michels, L., Tiefenbeck, V., & Maier, C. (2024). Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting. *Information Systems Journal, 34*(2), 384–414. https://doi.org/10.1111/isj.12482

Ötting, S. K., & Maier, G. W. (2018). The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behavior, 89*, 27–39. https://doi.org/10.1016/j.chb.2018.07.022

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Savitri, T. A., Buchori, I., & Supratikta, H. (2024). Exploring the role of Human Resources Information System in employee performance management: A systematic literature review. *Indonesian Development of Economics and Administration Journal, 3*(1), 55–64. https://doi.org/10.70001/idea.v3i1.211

Sumaryono, S. (2023). Transformasi e-HRM (Electronic Human Resource Management) dalam Pengembangan Kualitas SDM (Sumber Daya Manusia) Berbasis Teknologi Informasi Terhadap Kinerja Karyawan. *Jurnal Ekonomi Manajemen Sistem Informasi, 5*(1), 50–63. https://doi.org/10.31933/jemsi.v5i1.1680

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42. https://doi.org/10.1177/0008125619867910

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's Alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tong, S., Jia, N., Luo, X., & Fang, Z. (2021). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal, 42*(9), 1600–1631. https://doi.org/10.1002/smj.3322

Widyastuti, T., Ramdani, Z., Alim, S., & Harlianty, R. A. (2024). The Indonesian version of Individual Work Performance Questionnaire (IWPQ): Structure factor and criterion validity. *Psychological Research and Intervention, 7*(1), 1–16. https://doi.org/10.21831/pri.v7i1.76457

Yu, L., & Li, Y. (2022). Artificial intelligence decision-making transparency and employees' trust: The parallel multiple mediating effect of effectiveness and discomfort. *Behavioral Sciences, 12*(5), 127. https://doi.org/10.3390/bs12050127

How to cite this article: Irvan, M., Syahroni, B., & Yudi, M. (2026). Transparency, Algorithmic Fairness, and Employee Performance in AI-Based HR: The Moderating Role of Trust in Management. Indonesian Journal Economic Review (IJER), 6(1), 101-115. https://doi.org/10.59431/ijer.v6i1.729