

When Computational Fairness Is Not Enough: Explaining the Gap Between Algorithmic Fairness and Organizational Justice from a Business Ethics Perspective

Corresponding Author :

Seyed Jalalaldin Hosseini

Trainer, Technical and Vocational Training Organization (TVTO), Iran

Email: jalal602003@gmail.com

ORCID: 0009-0000-3481-5505

Compliance with Ethical Standards

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Research Involving Human Participants and/or Animals

This article is a conceptual study based on the analysis of previously published literature. It does not involve any studies with human participants or animals performed by the author.

Informed Consent

Informed consent was not required as no primary data were collected and no human participants were involved in this study.

When Computational Fairness Is Not Enough: Explaining the Gap Between Algorithmic Fairness and Organizational Justice from a Business Ethics Perspective

Abstract

Artificial intelligence (AI) is transforming organizational decision-making by improving efficiency, consistency, and scalability in human resource management. However, growing evidence suggests that algorithmically fair decisions do not necessarily lead to employees' perceptions of organizational justice. While existing research has extensively examined algorithmic fairness from computational and technical perspectives, limited attention has been paid to the ethical mechanisms that explain the gap between algorithmic fairness and organizational justice. Addressing this gap, this conceptual study develops a business ethics framework integrating organizational justice theory, ethical AI, algorithmic management, and organizational support theory. The proposed model argues that organizational justice is shaped not only by the fairness of algorithms but also by transparency, explainability, human oversight, accountability, ethical governance, and organizational trust. These mechanisms explain why technically fair AI systems may still generate perceptions of injustice when ethical and managerial processes are inadequate. The study extends organizational justice theory to the context of AI-enabled decision-making by distinguishing computational fairness from employees' justice perceptions. It contributes to the business ethics literature by providing a comprehensive conceptual framework for responsible AI governance and offers practical guidance for managers seeking to implement AI systems that are not only technically accurate but also ethically legitimate, trustworthy, and socially accepted within organizations.

Keywords: Algorithmic Fairness; Organizational Justice; Artificial Intelligence; Business Ethics; Algorithmic Management; Ethical AI; Human Resource Management.

Introduction

Beyond concerns regarding algorithmic bias, an increasing body of research argues that the ethical implications of AI in organizations cannot be understood solely through computational properties. Algorithmic systems are embedded within organizational structures, managerial practices, and institutional values that collectively shape employees' experiences of fairness (Meijerink et al., 2021; Bankins et al., 2024). Consequently, identical algorithms may generate substantially different justice perceptions depending on how organizations govern their implementation, communicate decisions, and preserve employees' dignity throughout the decision-making process.

This observation highlights an important conceptual distinction that remains insufficiently explored in the literature. Computational fairness represents an attribute of the algorithm itself, whereas organizational justice represents a psychological and ethical evaluation made by organizational members. The former is rooted primarily in mathematical optimization, statistical parity, and technical performance; the latter emerges from employees' interpretations of organizational motives, decision procedures, interpersonal treatment, and moral legitimacy (Colquitt, 2001; Goldman & Cropanzano, 2015). Treating these constructs as interchangeable risks overlooking the fundamentally social nature of organizational justice.

The distinction becomes even more significant when organizations increasingly delegate managerial authority to AI systems. Algorithmic management does not merely automate administrative tasks; it redistributes power within organizations by embedding managerial discretion into computational infrastructures (Kellogg et al., 2020; Parent-Rochelleau & Parker, 2022). Decisions that were once explained, negotiated, or reconsidered through interpersonal interactions may now be generated automatically, often with limited opportunities for employee participation or appeal. Such transformations alter not only how decisions are made but also how employees construct trust, legitimacy, and justice within the organization.

From an ethical standpoint, this shift raises a fundamental question: **Can an organizational decision be ethically fair simply because the algorithm producing it is computationally fair?**

Existing literature provides only partial answers. Research on explainable AI emphasizes transparency and interpretability (Burrell, 2016; de Fine Licht & de Fine Licht, 2020), while studies on AI ethics highlight accountability, privacy, and discrimination (Robert et al., 2020; Hunkenschroer & Luetge, 2022). Organizational justice scholars, meanwhile, continue to demonstrate that employees' reactions depend heavily

on procedural consistency, respectful treatment, adequate explanations, and opportunities to express their views (Colquitt et al., 2001). However, these research streams rarely converge into a unified ethical framework explaining how computational characteristics become—or fail to become—perceived organizational justice. The absence of such integration has practical consequences. Organizations frequently invest considerable resources in developing technically fair AI systems while paying comparatively less attention to employees' ethical perceptions of those systems. As a result, organizations may achieve compliance with technical fairness standards yet still experience declining trust, resistance to AI adoption, reduced organizational commitment, and diminished perceptions of legitimacy (Lee, 2018; Höddinghaus et al., 2021; Suseno et al., 2022). This suggests that ethical success in AI implementation depends not only on algorithm design but also on organizational governance.

Accordingly, we propose that the relationship between algorithmic fairness and organizational justice is neither direct nor automatic. Instead, it is shaped by ethical and organizational governance conditions that determine how employees interpret AI-supported decisions. Transparency, accountability, explainability, opportunities for voice, respect for human dignity, and organizational responsibility function as governance conditions that strengthen or weaken the extent to which computational fairness translates into perceptions of organizational justice.. Without these ethical conditions, even statistically unbiased algorithms may fail to achieve moral legitimacy inside organizations.

Based on this argument, this article develops a conceptual framework positioning **business ethics as the missing link** between computational fairness and organizational justice. Rather than viewing fairness exclusively as a technical property of AI systems, we conceptualize fairness as a multidimensional ethical phenomenon emerging from the interaction between algorithmic design, organizational governance, and employees' moral evaluations. This perspective extends organizational justice theory into AI-enabled workplaces while simultaneously enriching business ethics scholarship by demonstrating that ethical legitimacy cannot be delegated to algorithms alone.

Ultimately, this article argues that organizations should move beyond the question of "Is the algorithm fair?" toward the more comprehensive ethical question of "Is the organization fair when governing algorithmic decisions?" This shift redefines fairness from a computational objective into an organizational responsibility and provides a new theoretical foundation for responsible AI governance in contemporary organizations.

Although recent studies acknowledge that ethical governance and human oversight influence employee responses to AI-enabled decision-making (Robert et al., 2020; Meijerink et al., 2021; Bankins et al., 2024), existing research generally assumes a direct relationship between computational fairness and organizational justice. The present study challenges this assumption by arguing that computational fairness does not automatically translate into employees' justice perceptions. Instead, the strength of this relationship depends on the ethical governance conditions surrounding the decision. This study conceptualizes this moderating role as the organizational justice translation function, describing how governance conditions determine whether AI-generated decisions are interpreted, contextualized, and legitimized as organizationally fair.

2. Theoretical Background

2.1. Algorithmic Fairness: From Technical Fairness to Ethical Decision-Making

The rapid integration of artificial intelligence (AI) into organizational decision-making has fundamentally transformed human resource management by enabling automated recruitment, employee evaluation, promotion, and performance management. AI-driven systems promise greater consistency, efficiency, and objectivity than traditional human judgment, leading many organizations to rely increasingly on algorithmic decision-making in critical employment processes (Budhwar et al., 2022; Vrontis et al., 2022; Bankins et al., 2024).

Although algorithmic fairness has received considerable attention in computer science, contemporary AI ethics research argues that fairness represents only one component of responsible artificial intelligence. Ethical AI systems should simultaneously satisfy broader principles including transparency, accountability, explicability, human autonomy, and the promotion of societal well-being. Consequently, technical fairness alone cannot ensure that AI-supported organizational decisions will be perceived as ethically legitimate by employees because ethical evaluation encompasses a wider normative framework than computational optimization (Floridi & Cowls, 2019).

Ethical concerns surrounding algorithmic decision-making extend beyond statistical bias to include opacity, responsibility gaps, accountability, contestability, and the social consequences of automated decisions. Even mathematically fair algorithms may generate ethically problematic outcomes when individuals are unable to understand, challenge, or appeal algorithmic decisions. This broader ethical perspective reinforces the distinction between computational fairness and organizational justice developed in the present study (Mittelstadt et al., 2016).

Within computer science and data analytics, the concept of algorithmic fairness has primarily been developed as a technical objective aimed at minimizing systematic bias across demographic groups. Various computational fairness metrics—including demographic parity, equal opportunity, and equalized odds—seek to ensure that algorithmic outputs do not disproportionately disadvantage protected populations (Acikgoz et al., 2020; Köchling & Wehner, 2020; Robert et al., 2020). Considerable advances have therefore focused on improving data quality, mitigating algorithmic bias, and optimizing predictive accuracy.

Despite these technical developments, scholars increasingly argue that computational fairness represents only one dimension of fairness in organizational contexts. Algorithmic systems frequently operate as "black boxes," making their decision logic inaccessible to employees and managers alike (Burrell, 2016). Even when algorithms satisfy statistical fairness criteria, their opacity may prevent individuals from understanding why particular decisions have been made, thereby undermining confidence in organizational decision-making (de Fine Licht & de Fine Licht, 2020).

Moreover, organizational decisions are evaluated not only by their statistical outcomes but also by the legitimacy of the decision-making process. Employees frequently judge AI-assisted decisions according to whether they are transparent, understandable, contestable, and respectful of individual dignity rather than solely on whether the underlying algorithm is mathematically unbiased (Hunkenschroer & Luetge, 2022). Consequently, computational fairness should be regarded as a necessary—but not sufficient—condition for ethical organizational decision-making.

It is important to note that computational fairness is not itself a value-neutral or singular construct. The selection among competing fairness metrics—such as demographic parity, equal opportunity, or calibration—reflects normative assumptions about which form of equality is prioritized, and these metrics are frequently mutually incompatible (Kleinberg et al., 2017). Consequently, even the technical starting point of "computational fairness" already embeds contestable normative choices, a point that further reinforces the need to distinguish algorithmic fairness from organizational justice rather than treating the former as an objective baseline against which the latter is measured.

2.2. Distinguishing Algorithmic Fairness from Organizational Justice

A fundamental premise of this study is that algorithmic fairness and organizational justice represent conceptually distinct constructs despite their frequent conflation within contemporary AI and HRM research. Existing studies often assume that improving computational fairness automatically produces fair organizational outcomes. However, this assumption overlooks the fundamental distinction between technical optimization and employees' social and ethical evaluations of organizational decisions (Acikgoz et al., 2020; Newman et al., 2020).

Algorithmic fairness refers to the technical properties of AI systems designed to reduce systematic bias and ensure equitable treatment across individuals or demographic groups. Within computer science, fairness is commonly operationalized through statistical criteria such as demographic parity, equal opportunity, equalized odds, calibration, and predictive consistency. These computational metrics evaluate whether algorithms generate mathematically unbiased predictions under predefined conditions (Burrell, 2016; de Fine Licht & de Fine Licht, 2020; Robert et al., 2020).

Organizational justice, in contrast, is grounded in organizational behavior and business ethics and concerns employees' subjective judgments regarding the fairness of organizational decisions, managerial actions, and interpersonal treatment. According to Colquitt (2001), justice perceptions encompass distributive, procedural, interpersonal, and informational dimensions, each reflecting different aspects of fairness experienced by organizational members. Unlike algorithmic fairness, organizational justice extends beyond statistical consistency to include ethical legitimacy, transparency, respect, accountability, and opportunities for employee voice (Colquitt, 2001; Goldman & Cropanzano, 2015).

Consequently, computational fairness should be viewed as a necessary—but not sufficient—condition for organizational justice. An AI system may satisfy multiple statistical fairness criteria while simultaneously generating perceptions of injustice if employees cannot understand decision logic, question outcomes, appeal unfavorable decisions, or perceive meaningful human oversight. Justice therefore emerges not solely from algorithmic performance but from the broader organizational context in which AI-enabled decisions are governed and implemented (Lee, 2018; Hunkenschroer & Luetge, 2022; Bankins et al., 2024).

Building upon this distinction, the present study conceptualizes ethical organizational governance as the condition that determines whether technically fair algorithms translate into employees' perceptions of

organizational justice. This perspective shifts attention from computational optimization toward responsible organizational implementation, thereby providing a business ethics explanation for why mathematically fair AI systems do not necessarily produce ethically legitimate organizational outcomes.

This framework further distinguishes three analytically separate levels: computational assessment (whether an algorithm meets a defined statistical criterion), perceived organizational justice (employees' subjective evaluation of fairness), and normative ethical legitimacy (whether a decision-making process is defensible according to independent ethical standards). These three levels do not necessarily converge. A decision may be computationally fair yet perceived as unjust; conversely, a decision may be perceived as fair by employees while remaining ethically indefensible—for instance, when employees lack the information needed to recognize a procedural flaw. Positive employee perceptions therefore cannot, by themselves, establish that a system is ethically legitimate; they are one input to, but not a substitute for, normative evaluation.

A related conceptual risk concerns the potential overlap between ethical governance mechanisms and the procedural and informational dimensions of organizational justice as defined by Colquitt (2001). To avoid tautology, this study distinguishes governance as an organizational-level practice (e.g., the existence of an appeal process, the publication of decision criteria) from justice as an individual-level perception (e.g., an employee's subjective judgment that the appeal process was fair). Governance mechanisms are necessary but not sufficient conditions for procedural and informational justice perceptions: an organization may implement transparent policies that employees nonetheless perceive as inadequate, superficial, or symbolic. The framework therefore treats governance and justice as related but empirically separable constructs operating at different levels of analysis (organizational practice versus individual perception).

International AI governance initiatives increasingly emphasize transparency, accountability, human oversight, explainability, and fairness as foundational principles for trustworthy AI deployment. Comparative analyses of global AI ethics guidelines demonstrate remarkable convergence around these governance principles, suggesting that organizational implementation of AI requires ethical governance structures in addition to technically robust algorithms (Jobin et al., 2019).

From an organizational ethics perspective, responsible AI implementation requires governance mechanisms capable of balancing technological innovation with ethical accountability. Human oversight, explainability,

organizational responsibility, and continuous ethical evaluation constitute essential safeguards that enable organizations to transform technically accurate AI systems into ethically acceptable decision-making processes (Floridi et al., 2018).

2.3. Organizational Justice Theory

Organizational justice provides one of the most influential theoretical frameworks for understanding how employees evaluate fairness within organizations. Rather than focusing exclusively on decision outcomes, organizational justice theory argues that employees develop fairness judgments through their evaluation of organizational procedures, interpersonal treatment, and the quality of information provided during decision-making processes (Colquitt, 2001; Colquitt et al., 2001).

The literature commonly conceptualizes organizational justice as comprising four complementary dimensions. Distributive justice concerns the perceived fairness of outcomes and resource allocation (Leventhal, 1980). Procedural justice reflects the perceived fairness, consistency, impartiality, and ethicality of decision-making procedures. Interpersonal justice refers to whether individuals are treated with dignity, politeness, and respect during organizational interactions. Finally, informational justice concerns the adequacy, honesty, and transparency of explanations provided regarding organizational decisions (Colquitt, 2001).

Extensive empirical evidence demonstrates that perceptions of organizational justice influence numerous organizational outcomes, including organizational commitment, trust, job satisfaction, employee engagement, organizational citizenship behavior, and perceived organizational support (Colquitt et al., 2001; Eisenberger et al., 1986; Rhoades & Eisenberger, 2002; Kurtessis et al., 2017). Importantly, fairness perceptions emerge from employees' subjective interpretation of organizational practices rather than from objective characteristics of decision systems alone.

Goldman and Cropanzano (2015) further distinguish between fairness and justice, arguing that fairness reflects an individual's subjective perception, whereas justice represents a broader normative evaluation grounded in ethical principles. This distinction becomes particularly important when organizational decisions are delegated to artificial intelligence, where mathematically fair algorithms may nevertheless fail to generate perceptions of justice among organizational members.

2.4. The Gap Between Algorithmic Fairness and Organizational Justice

Although algorithmic fairness and organizational justice both address issues of equity, they originate from different disciplinary traditions and rely on fundamentally different assumptions. Algorithmic fairness primarily evaluates whether computational models produce statistically unbiased outcomes, whereas organizational justice examines how individuals psychologically experience organizational decision-making.

Emerging evidence suggests that these two forms of fairness do not necessarily coincide. Employees frequently express skepticism toward AI-generated decisions even when algorithms demonstrate higher predictive accuracy and lower statistical bias than human decision makers (Lee, 2018; Mirowska & Mesnet, 2022). Individuals often perceive algorithmic decisions as unfair because they lack meaningful explanations, opportunities for participation, mechanisms for appeal, or identifiable human accountability (Newman et al., 2020).

Algorithmic opacity further amplifies these concerns. When employees cannot understand how decisions are generated, they may question not only the validity of individual decisions but also the ethical legitimacy of the organization itself (Burrell, 2016; Paudyal & Wong, 2018). Likewise, excessive reliance on algorithmic management may reduce perceptions of autonomy, diminish trust, and weaken employees' sense of procedural justice (Kellogg et al., 2020; Parent-Rochelleau & Parker, 2022).

Research on algorithmic management similarly demonstrates that AI-driven control systems often reshape employment relationships by transferring managerial authority from human supervisors to automated systems (Duggan et al., 2020; Möhlmann et al., 2021; Meijerink et al., 2021). Under such circumstances, employees evaluate organizational fairness not merely according to computational outcomes but also according to whether organizations maintain transparency, accountability, and meaningful human oversight throughout the decision-making process.

Accordingly, a critical theoretical distinction emerges: an algorithm may satisfy computational fairness requirements while simultaneously failing to generate perceptions of organizational justice. This divergence represents a significant conceptual gap that existing research has only partially addressed.

2.5. A Business Ethics Perspective: Beyond Computational Fairness

From the perspective of business ethics, ethical organizational decision-making cannot be reduced to statistically unbiased algorithmic outputs. Ethical legitimacy requires that organizational decisions respect

fundamental principles such as transparency, accountability, explainability, human dignity, and responsible governance.

Recent scholarship in AI ethics increasingly argues that organizations deploying AI systems bear moral responsibilities extending beyond technical performance (Hunkenschroer & Luetge, 2022; Budhwar et al., 2022). Ethical AI therefore requires organizations to ensure that employees understand how important decisions are made, possess opportunities to question automated judgments, and retain meaningful human review where appropriate.

Business ethics scholars further emphasize that trust in AI-enabled organizations depends not only on algorithmic accuracy but also on employees' perceptions that organizational values, ethical norms, and human rights remain central throughout technology adoption (Bankins et al., 2024; Giermindl et al., 2022). Consequently, organizations should view AI not merely as a technological innovation but as an ethical governance challenge requiring responsible managerial oversight.

Building upon organizational justice theory and contemporary business ethics, this study argues that computational fairness alone cannot fully explain employees' justice perceptions. Instead, organizational justice represents a broader ethical construct shaped by transparency, procedural legitimacy, accountability, and respect for human dignity. Accordingly, this paper proposes that the discrepancy between computational fairness and organizational justice constitutes an ethical gap that deserves independent theoretical attention. By conceptualizing this ethical gap, the study contributes to the growing literature on responsible AI and business ethics while extending organizational justice theory into the emerging domain of AI-enabled human resource management.

3. Methodology

3.1 Research Design

This study adopts a conceptual research design to investigate the ethical relationship between algorithmic fairness and organizational justice in AI-supported organizational decision-making. Rather than empirically testing causal relationships, the study seeks to develop an integrated theoretical framework explaining why

technically fair algorithms may nevertheless fail to produce perceptions of organizational justice among employees.

Conceptual research is particularly appropriate when existing knowledge is fragmented across multiple disciplines and when new theoretical integration is required to explain emerging organizational phenomena. The rapid adoption of artificial intelligence in human resource management represents such a context, where insights from computer science, organizational behavior, business ethics, and human resource management remain insufficiently integrated.

Accordingly, this study aims to synthesize these diverse bodies of literature to construct a coherent ethical framework capable of explaining the conceptual gap between computational fairness and organizational justice.

3.2 Literature Selection Strategy

The theoretical framework was developed through a structured review of the contemporary literature on artificial intelligence, algorithmic management, organizational justice, and business ethics.

The literature selection emphasized peer-reviewed journal articles published in internationally recognized journals indexed in the Web of Science and Scopus databases. Particular attention was given to publications appearing in leading journals including *Journal of Business Ethics*, *Human Resource Management Review*, *The International Journal of Human Resource Management*, *Journal of Organizational Behavior*, *Academy of Management Annals*, *MIS Quarterly*, *European Journal of Information Systems*, and *Big Data & Society*.

Priority was assigned to studies published between 2015 and 2024 because this period reflects the rapid emergence of algorithmic management and responsible AI research. Classic foundational works on organizational justice and perceived organizational support were also incorporated where necessary to establish the theoretical foundations of the proposed framework.

The search strings used were combinations such as ("algorithmic fairness" OR "AI fairness") AND ("organizational justice" OR "procedural justice"), and ("algorithmic management") AND ("business ethics" OR "AI governance"). Searches were performed in Web of Science, Scopus, and Google Scholar, covering publications from 2015 to 2024. An initial search yielded approximately 140 records across the three databases. After removing duplicates (18 records) and screening titles and abstracts for direct relevance to the intersection

of algorithmic decision-making and organizational or ethical evaluation, 62 full-text articles were retained for synthesis. Articles were excluded if they addressed algorithmic fairness purely as a technical/statistical problem without organizational or ethical implications, or if they addressed organizational justice without reference to algorithmic or AI-based decision-making. Non-peer-reviewed sources were excluded, with the exception of official governance documents (e.g., OECD, European Commission, European Union).

3.3 Analytical Approach

The study employs theory synthesis as its primary analytical approach. Theory synthesis seeks to integrate complementary theoretical perspectives into a unified conceptual explanation of a phenomenon that cannot be adequately explained by any single theoretical lens.

The analysis proceeded in four stages.

First, the principal concepts of algorithmic fairness were identified from the AI ethics and algorithmic management literature.

Second, the core assumptions of organizational justice theory were examined, with particular emphasis on distributive, procedural, interpersonal, and informational justice.

Third, these two streams of literature were systematically compared to identify conceptual inconsistencies, overlapping assumptions, and unresolved theoretical tensions.

Finally, insights from business ethics were incorporated to explain why computational fairness alone cannot ensure ethical legitimacy within organizational decision-making, resulting in the development of an integrated conceptual framework.

3.4 Theoretical Contribution

The contribution of this study is theory-building rather than theory-testing.

Instead of proposing new statistical evidence, the paper introduces a conceptual model demonstrating that algorithmic fairness constitutes only one component of ethical organizational decision-making. Organizational justice is argued to represent a broader construct encompassing transparency, explainability, accountability, procedural legitimacy, and respect for human dignity. By integrating organizational justice theory with business ethics and contemporary AI governance literature, the study develops the concept of an ethical gap between

computational fairness and employees' justice perceptions. This framework contributes to the growing literature on responsible AI while extending organizational justice theory to AI-enabled organizational contexts.

4. Conceptual Framework

This framework operates across three levels of analysis: the system level (computational fairness metrics), the organizational level (governance practices), and the individual level (employee perceptions of justice and trust). Organizational outcomes (e.g., turnover, legitimacy) represent an aggregate consequence of individual-level perceptions and are treated as a distal, rather than proximal, outcome of the framework's core mechanisms.

The conceptual framework proposed in this study explains why computational fairness alone is insufficient to generate perceptions of organizational justice in AI-enabled organizations. Building upon the business ethics literature, the model distinguishes between algorithmic fairness, which refers to the technical properties of AI decision-making systems, and organizational justice, which reflects employees' ethical, psychological, and social evaluations of organizational decisions (Robert et al., 2020; Bankins et al., 2024).

The framework argues that algorithmic fairness is a necessary but insufficient condition for achieving organizational justice. From a computational perspective, AI systems may satisfy recognized fairness criteria—including bias mitigation, consistency, transparency, explainability, accountability, and predictive accuracy—yet employees may still perceive AI-supported decisions as unfair because justice evaluations extend beyond statistical optimization and incorporate moral legitimacy, interpersonal treatment, procedural integrity, and organizational responsibility (Burrell, 2016; de Fine Licht & de Fine Licht, 2020; Köchling & Wehner, 2020; Robert et al., 2020).

Consistent with organizational justice theory, employees evaluate AI-assisted decisions through four complementary dimensions of justice: distributive justice, procedural justice, interpersonal justice, and informational justice (Colquitt, 2001; Colquitt et al., 2001). While algorithmic fairness may positively influence these justice perceptions, this influence is neither automatic nor deterministic. Employees interpret algorithmic decisions within a broader organizational context that includes managerial accountability, transparency of decision-making processes, opportunities for employee voice and appeal, ethical leadership, and meaningful human oversight (Newman et al., 2020; Hunkenschroer & Luetge, 2022; Bankins et al., 2024).

Rather than treating ethical governance as a mediating pathway through which algorithmic fairness produces organizational justice—which would imply that governance is itself a consequence of algorithmic fairness—this framework treats ethical governance as a moderating condition. Ethical governance does not arise from computational fairness; rather, it determines the extent to which technical properties of an algorithm become visible, contestable, and correctable to employees. Accordingly, the strength of the relationship between algorithmic fairness and organizational justice depends on the level of ethical governance mechanisms in place—including responsible AI policies, explainability practices, accountability structures, ethical oversight, employee participation, and human review of critical decisions (Meijerink et al., 2021; Robert et al., 2020; Malik et al., 2022; Bankins et al., 2024)—rather than governance being caused by algorithmic fairness itself.

The framework further proposes that stronger perceptions of organizational justice generate a series of favorable organizational outcomes, including greater trust in AI-assisted decision-making, stronger organizational legitimacy, higher employee acceptance of AI technologies, greater organizational commitment, and ultimately more responsible and sustainable AI implementation (Lee, 2018; Acikgoz et al., 2020; Suseno et al., 2022; Vrontis et al., 2022). Conversely, organizations that emphasize computational fairness while neglecting ethical governance are likely to encounter employee distrust, resistance to AI-enabled HR practices, and diminished perceptions of justice despite technically unbiased algorithms (Köchling & Wehner, 2020; Duggan et al., 2020; Bankins et al., 2024).

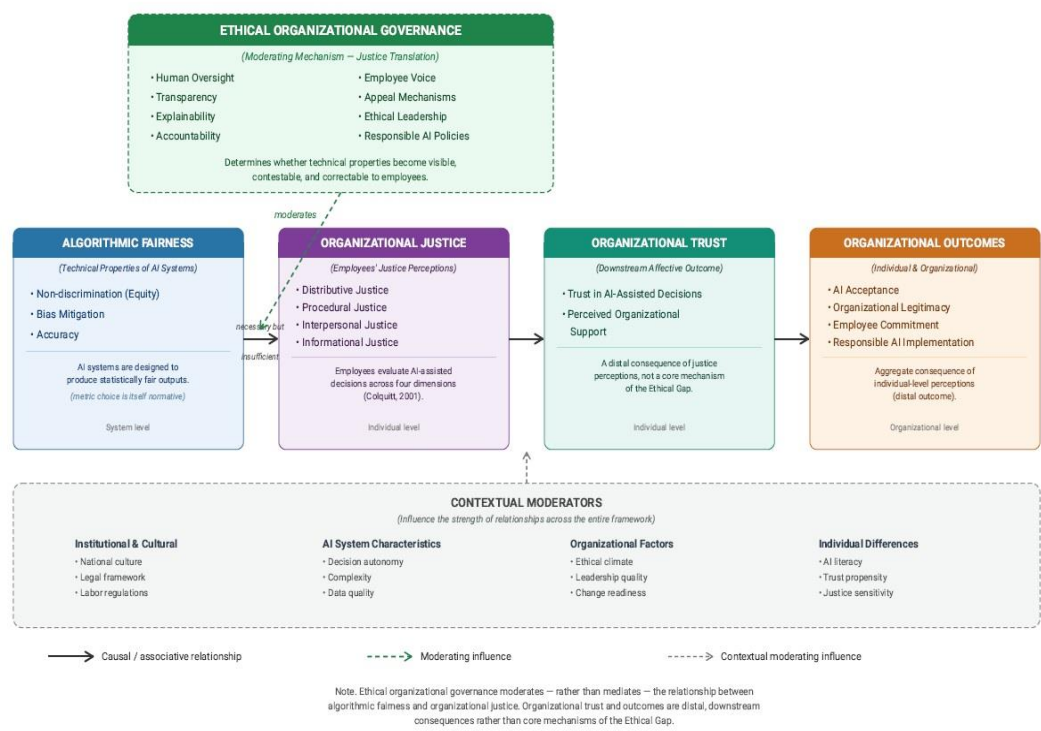
Organizational trust and perceived organizational support (Eisenberger et al., 1986; Kurtessis et al., 2017) are positioned in this framework as downstream, affective outcomes of organizational justice perceptions rather than as independent mechanisms within the core Ethical Gap model. They are included to indicate the framework's boundary conditions for future empirical testing, not as constructs requiring separate theoretical development here.

Finally, the framework recognizes that the relationship between algorithmic fairness and organizational justice is shaped by several contextual conditions. Cultural and institutional environments, characteristics of AI systems (e.g., autonomy, transparency, and complexity), organizational climate, ethical culture, and individual differences such as AI literacy, trust propensity, and justice sensitivity may strengthen or weaken employees' justice perceptions. These contextual moderators acknowledge that perceptions of fairness emerge not only

from technological design but also from the broader ethical and organizational environment in which AI systems operate (Parent-Rochelleau & Parker, 2022; Giermindl et al., 2022; Bankins et al., 2024).

The proposed conceptual framework is presented in Figure 1.

Figure 1. Conceptual framework explaining the gap between algorithmic fairness and organizational justice (revised model: ethical governance as moderator)



5. Conceptual Model and Proposition Development

5.1 From Computational Fairness to Ethical Legitimacy

Existing research on algorithmic fairness has primarily concentrated on improving computational performance through statistical fairness metrics, bias mitigation techniques, explainable artificial intelligence (XAI), and transparency mechanisms (Burrell, 2016; Köchling & Wehner, 2020; Robert et al., 2020). These advances have significantly enhanced the technical reliability of algorithmic decision-making in organizational settings. However, technical fairness does not necessarily translate into ethical legitimacy. Employees evaluate

organizational decisions through broader social and moral standards rather than computational criteria alone. Organizational justice theory suggests that individuals assess decisions according to the fairness of procedures, the quality of interpersonal treatment, the adequacy of explanations, and the perceived intentions of decision makers (Colquitt, 2001; Colquitt et al., 2001).

Consequently, an algorithm may satisfy mathematical fairness constraints while simultaneously generating perceptions of injustice because employees remain unable to understand, challenge, or morally evaluate algorithmic decisions. This distinction represents the central theoretical tension addressed in this study.

5.2 Bridging Algorithmic Fairness and Organizational Justice

This study argues that algorithmic fairness represents only one dimension of ethical organizational decision-making. Business ethics requires organizational decisions to satisfy broader principles including accountability, transparency, explainability, respect for human dignity, and moral responsibility (de Fine Licht & de Fine Licht, 2020; Hunkenschroer & Luetge, 2022). From this perspective, organizational justice functions as the ethical mechanism through which employees interpret algorithmic decisions. Therefore, justice perceptions are influenced not only by computational outputs but also by organizational governance surrounding AI implementation. This reasoning explains why organizations employing technically unbiased algorithms may nevertheless experience employee resistance, declining trust, lower organizational commitment, and reduced acceptance of AI-supported decisions.

5.3 The Ethical Gap

Building upon organizational justice theory and contemporary business ethics literature, this study introduces the concept of the Ethical Gap.

The Ethical Gap refers to the discrepancy between objectively fair algorithmic outputs and employees' subjective perceptions of organizational justice.

This gap emerges whenever algorithmic decisions satisfy computational fairness requirements while failing to meet ethical expectations regarding transparency, procedural legitimacy, explainability, accountability, and respect for employees.

Accordingly, ethical legitimacy depends not only on algorithmic design but also on organizational governance practices that shape employees' justice perceptions.

The Ethical Gap proposed here should be distinguished from related constructs such as algorithmic reductionism (Newman et al., 2020), which describes how the removal of contextual and qualitative information from decision-making can itself generate perceptions of injustice, independent of governance conditions. Whereas algorithmic reductionism explains injustice as a property of the algorithm's informational scope, the Ethical Gap concept advanced in this study locates the discrepancy at the organizational level: it emerges from the absence of governance mechanisms surrounding an otherwise technically fair algorithm, rather than from the algorithm's internal logic alone. The Ethical Gap is therefore a governance-level construct, complementary to but conceptually distinct from algorithm-level explanations of perceived unfairness.

5.3.1 A Configurational View of the Ethical Gap

To operationalize the Ethical Gap, this framework proposes a 2×2 configurational typology crossing computational fairness (high/low) with perceived organizational justice (high/low):

Aligned Legitimacy (high computational fairness, high perceived justice): the condition organizations aim for, where technical and social evaluations converge.

Legitimacy Deficit (high computational fairness, low perceived justice): the central case addressed by this study, in which technically sound algorithms fail to generate employee trust due to inadequate governance.

False Legitimacy (low computational fairness, high perceived justice): a concerning configuration in which employees perceive fairness despite underlying statistical bias, often due to opacity or insufficient contestability—here, positive perceptions mask, rather than reflect, injustice.

Manifest Injustice (low computational fairness, low perceived justice): both technical and perceptual failure, typically the most visible and reputationally damaging condition.

This typology clarifies that the Ethical Gap specifically denotes the Legitimacy Deficit quadrant, while also drawing attention to False Legitimacy as an underexamined risk in which the absence of employee complaints does not indicate ethical soundness.

5.4 Conceptual Propositions

Based on the proposed theoretical framework, the following propositions are advanced.

Proposition 1

The positive relationship between algorithmic fairness and organizational justice perceptions is moderated by ethical governance, such that this relationship is stronger under high ethical governance than under low ethical governance.

Proposition 2

The Ethical Gap (the discrepancy between computational fairness and perceived organizational justice) widens as AI decision autonomy increases, independent of the algorithm's statistical fairness properties.

Proposition 3

Perceived organizational justice under conditions of low computational fairness but high procedural transparency (the "False Legitimacy" configuration) is positively associated with organizational trust in the short term, but negatively associated with organizational trust once underlying bias becomes visible to employees.

Proposition 4

The negative relationship between the Ethical Gap and organizational trust is stronger for high-stakes decisions (e.g., dismissal, promotion) than for low-stakes decisions (e.g., scheduling, routine feedback).

Proposition 5

Organizational governance mechanisms (e.g., appeal procedures) are more strongly associated with perceived procedural justice when employees attribute the governance mechanism to genuine organizational commitment rather than to symbolic or reputation-driven motives.

5.5 Proposed Conceptual Framework

The framework proposes that algorithmic fairness alone cannot directly produce organizational justice. Instead, organizational justice emerges through ethical governance mechanisms including transparency, explainability, accountability, procedural legitimacy, and respect for human dignity. These ethical mechanisms reduce the Ethical Gap and ultimately enhance employees' trust, commitment, and acceptance of AI-supported organizational decisions.

6. Discussion

6.1. Reframing Organizational Justice in the Age of Artificial Intelligence

The rapid integration of artificial intelligence into organizational decision-making has fundamentally transformed how organizations recruit, evaluate, promote, and manage employees. While recent research has devoted considerable attention to improving algorithmic fairness through bias mitigation, explainability, and transparency, the prevailing literature continues to treat fairness predominantly as a computational property of algorithms (Acikgoz et al., 2020; Köchling & Wehner, 2020; Robert et al., 2020). This perspective, although technically valuable, overlooks a fundamental ethical reality: organizational decisions are ultimately interpreted by human beings rather than machines.

The findings of the present study align with the Trustworthy AI framework proposed by the European Commission, which argues that fairness should operate alongside transparency, accountability, human agency, technical robustness, and societal well-being. These principles collectively support the argument that organizational justice cannot emerge solely from computational fairness but depends upon ethically governed AI implementation. (European Commission. 2019)

By distinguishing algorithmic fairness from organizational justice, this study challenges the implicit assumption that technically unbiased AI systems automatically produce ethically legitimate organizational outcomes.

More recently, the European Union's Artificial Intelligence Act (2024) has begun to codify several of the governance principles discussed in this framework—including transparency obligations, human oversight requirements, and contestability rights—for high-risk AI systems, which explicitly include employment-related decision-making. This regulatory development lends practical urgency to the theoretical distinction advanced here: organizations will increasingly be required, not merely encouraged, to treat ethical governance as a formal condition for the lawful deployment of AI in HR contexts.

6.2 Rival Explanations

Several alternative explanations could account for the divergence between algorithmic fairness and organizational justice without invoking ethical governance as a distinct construct. Outcome favorability suggests that employees judge fairness based on whether the decision benefited them personally, regardless of process (Newman et al., 2020). Algorithm aversion (Dietvorst et al., 2015) proposes that employees distrust

algorithmic decisions more than equivalent human decisions independent of governance quality. Prior organizational trust may condition how any AI-supported decision is received, such that low baseline trust produces injustice perceptions regardless of governance. Task appropriateness suggests that perceived injustice arises specifically when AI is applied to decisions employees consider inherently requiring human judgment (e.g., termination) rather than from governance failures per se. Power relations frameworks (Kellogg et al., 2020) locate injustice in the redistribution of managerial authority itself, independent of transparency measures. Symbolic governance describes cases in which formal governance structures exist but are perceived by employees as performative rather than substantive, producing continued injustice despite nominal compliance with transparency and appeal requirements. Finally, reverse causality is possible: employees who already distrust an organization may retrospectively interpret computationally fair decisions as unjust, rather than governance quality causing justice perceptions.

The framework advanced in this study does not claim to supersede these explanations but proposes that ethical governance operates as a distinct, testable moderating condition that can be examined alongside—rather than instead of—these rival accounts in future empirical research.

6.3. The Ethical Gap as a New Conceptual Lens

The principal theoretical contribution of this article is the introduction of the concept of the Ethical Gap.

6.4. Theoretical Contributions

This article contributes to three complementary streams of literature.

First, it extends organizational justice theory by demonstrating that justice perceptions in AI-supported workplaces depend not only on distributive and procedural outcomes but also on employees' ethical evaluation of algorithmic governance.

Second, it contributes to the rapidly expanding literature on algorithmic management and AI-enabled HRM by integrating organizational justice theory with business ethics. Existing research has largely examined algorithmic systems from technological, managerial, or legal perspectives (Meijerink et al., 2021; Budhwar et al., 2022; Bankins et al., 2024). The present framework positions ethical legitimacy as the missing mechanism connecting computational systems with employee reactions.

Ethical governance conditions the strength of the relationship between technically fair algorithmic outputs and organizational justice evaluations. Ethical governance performs a justice translation function, determining the extent to which computational assessments are interpreted, contextualized, legitimized, and transformed into organizational justice perceptions. Accordingly, this study conceptualizes this moderating function as the Organizational Justice Translation Mechanism, providing a novel ethical explanation for the gap between computational fairness and organizational justice.

Third, the proposed framework enriches business ethics scholarship by emphasizing that ethical AI cannot be achieved solely through technical design. Responsible AI implementation requires continuous organizational commitment to transparency, accountability, contestability, human oversight, and respect for employee autonomy.

6.5. Practical Implications

The proposed framework carries important implications for organizational leaders, HR professionals, AI developers, and policymakers.

Organizations should recognize that investing exclusively in technically fair algorithms is insufficient for establishing employee trust. Ethical governance mechanisms—including transparent communication, explainable decision processes, accessible appeal procedures, and meaningful human oversight—should accompany every stage of AI implementation.

Human resource departments should preserve opportunities for managerial discretion in sensitive employment decisions rather than transferring complete authority to automated systems. Employees should receive understandable explanations regarding algorithmic decisions and possess accessible mechanisms to challenge or review automated outcomes.

Similarly, AI developers should expand evaluation criteria beyond predictive performance and fairness metrics by incorporating ethical principles into algorithm design. Integrating organizational justice considerations during system development can substantially improve employee acceptance and long-term organizational legitimacy.

Organizations implementing AI-supported HR systems should align internal governance mechanisms with internationally recognized AI governance principles. The OECD AI Principles recommend transparency, accountability, human-centered values, and robust governance as prerequisites for trustworthy AI, providing organizations with practical guidance for transforming computational fairness into sustainable organizational justice (OECD, 2019).

7. Conclusion

The rapid integration of artificial intelligence into organizational decision-making has transformed human resource management and intensified scholarly attention to the concept of algorithmic fairness (Budhwar et al., 2022; Vrontis et al., 2022; Bankins et al., 2024). However, this article argues that computational fairness, while necessary, is not sufficient to ensure organizational justice. Even when AI systems satisfy technical fairness criteria and reduce measurable forms of algorithmic bias, employees continue to evaluate organizational decisions according to broader ethical expectations involving transparency, accountability, procedural legitimacy, opportunities for voice, and respect for human dignity (Colquitt, 2001; Goldman & Cropanzano, 2015; de Fine Licht & de Fine Licht, 2020).

Building upon organizational justice theory, business ethics, and the emerging literature on algorithmic management, this conceptual paper distinguishes algorithmic fairness from organizational justice and demonstrates that these concepts are theoretically, psychologically, and ethically distinct (Acikgoz et al., 2020; Newman et al., 2020; Köchling & Wehner, 2020). The article proposes that organizations may experience an Ethical Gap whenever technically fair algorithmic decisions fail to satisfy employees' ethical expectations regarding how decisions are explained, governed, challenged, and justified. This gap represents the central theoretical contribution of the study and provides a new lens for understanding why perceptions of injustice persist despite advances in AI fairness technologies.

Ultimately, this article advances a fundamental proposition for both scholars and practitioners: fair algorithms do not necessarily create fair organizations. Organizations become genuinely just not because their algorithms satisfy mathematical fairness metrics, but because their governance systems ensure that AI-assisted decisions are transparent, accountable, contestable, and ethically legitimate. Closing the gap between computational fairness and organizational justice therefore represents one of the defining ethical challenges for organizations

in the age of artificial intelligence and an essential prerequisite for responsible AI governance in the future of work (Robert et al., 2020; Bankins et al., 2024; Budhwar et al., 2022).

8.Limitations and Future Research

Despite its theoretical contributions, this study has several limitations that provide promising avenues for future research.

First, this research adopts a conceptual approach rather than an empirical design. Although conceptual analyses are essential for theory development in emerging fields such as algorithmic management and AI ethics, future studies should empirically validate the proposed framework across different organizational settings, industries, and cultural contexts (Bankins et al., 2024; Budhwar et al., 2022). Longitudinal and multi-level studies would be particularly valuable in examining how employees' justice perceptions evolve as organizations progressively integrate AI into HR decision-making.

Second, the present framework treats organizational justice primarily through the dimensions of distributive, procedural, interpersonal, and informational justice (Colquitt, 2001). Future research could extend this perspective by incorporating ethical constructs such as legitimacy, accountability, explainability, trust in AI, organizational transparency, and employee voice, thereby providing a more comprehensive understanding of ethical AI governance (de Fine Licht & de Fine Licht, 2020; Lee, 2018; Robert et al., 2020).

Third, although the study argues that computational fairness cannot fully substitute organizational justice, different AI systems employ varying algorithmic architectures, levels of transparency, and human oversight. Future investigations should therefore compare different forms of algorithmic decision-making—including recruitment algorithms, performance management systems, promotion algorithms, and people analytics—to determine whether justice perceptions differ across technological contexts (Cheng & Hackett, 2021; Meijerink et al., 2021; Giermindl et al., 2022).

Another limitation concerns contextual generalizability. Employee responses to AI-assisted decisions are likely to differ across institutional environments, regulatory systems, and national cultures. Comparative international studies could investigate how cultural values, legal frameworks, and organizational governance mechanisms influence the relationship between algorithmic fairness and organizational justice (Acikgoz et al., 2020; Vrontis et al., 2022).

Future research should also examine the psychological mechanisms through which employees evaluate AI-supported managerial decisions. Variables such as perceived organizational support, organizational trust, psychological safety, organizational identification, and commitment may function as mediators or moderators that strengthen or weaken justice perceptions (Eisenberger et al., 1986; Kurtessis et al., 2017; Meyer & Allen, 1991).

Finally, the rapid evolution of generative AI, autonomous decision systems, and human–AI collaboration suggests that organizational justice itself may require theoretical refinement. Rather than treating justice as a static organizational construct, future scholars may explore how ethical governance, human oversight, algorithmic explainability, and collaborative intelligence jointly shape what could be conceptualized as "responsible algorithmic justice" within AI-enabled organizations (Hunkenschroer & Luetge, 2022; Bankins et al., 2024; Robert et al., 2020).

References

- Acikgoz, Y., Davison, K., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, 28(4), 399–416.
<https://doi.org/10.1111/ijsa.12306>
- Bankins, S., Ocampo, A. C., 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>
- Budhwar, P., Malik, A., De Silva, M. T., & Thevisuthan, P. (2022). Artificial intelligence: Challenges and opportunities for international HRM: A review and research agenda. *The International Journal of Human Resource Management*, 33(6), 1065–1097. <https://doi.org/10.1080/09585192.2022.2035161>
- Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
- Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, 31(1), 100698. <https://doi.org/10.1016/j.hrmr.2019.100698>
- 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>
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & 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>
- de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making. *AI & Society*, 35(4), 917–926. <https://doi.org/10.1007/s00146-020-00960-w>
- Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal*, 30(1), 114–132. <https://doi.org/10.1111/1748-8583.12258>

- 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>
- Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology*, 71(3), 500–507. <https://doi.org/10.1037/0021-9010.71.3.500>
- European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Giermindl, L. M., Strich, F., Christ, O., Leicht-Deobald, U., & Redzepi, A. (2022). The dark sides of people analytics: Reviewing the perils for organisations and employees. *European Journal of Information Systems*, 31(3), 410–435. <https://doi.org/10.1080/0960085X.2021.1927213>
- Höddinghaus, M., Sondern, D., & Hertel, G. (2021). The automation of leadership functions: Would people trust decision algorithms? *Computers in Human Behavior*, 116, 106635. <https://doi.org/10.1016/j.chb.2020.106635>
- Hunkenschroer, A., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178(4), 977–1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Jabagi, N., Croteau, A.-M., & Audebrand, L. (2020). Perceived organizational support in the face of algorithmic management: A conceptual model. *Proceedings of the Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2020.489>
- 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 HR recruitment and HR development. *Business Research*, 13(3), 795–848. <https://doi.org/10.1007/s40685-020-00134-w>

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2017). Inherent trade-offs in the fair determination of risk scores. In 8th Innovations in Theoretical Computer Science Conference (ITCS 2017) (pp. 43:1–43:23).

Leibniz International Proceedings in Informatics. <https://doi.org/10.4230/LIPIcs.ITCS.2017.43>

Kurtessis, J. N., Eisenberger, R., Ford, M. T., Buffardi, L. C., Stewart, K. A., & Adis, C. S. (2017). Perceived organizational support: A meta-analytic evaluation of organizational support theory. *Journal of Management*, 43(6), 1854–1884. <https://doi.org/10.1177/0149206315575554>

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>

Malik, A., Budhwar, P., Patel, C., & Srikanth, N. (2022). May the bots be with you! Delivering HR cost-effectiveness and individualised employee experiences in an MNE. *The International Journal of Human Resource Management*, 33(6), 1148–1178.

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.

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.

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), 100838.

Rhoades, L., & Eisenberger, R. (2002). Perceived organizational support: A review of the literature. *Journal of Applied Psychology*, 87(4), 698–714.

Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human–Computer Interaction*, 35(5–6), 545–575.

Suseno, Y., Chang, C., Hudik, M., & Fang, E. (2022). Beliefs, anxiety and change readiness for artificial intelligence adoption among HR managers. *The International Journal of Human Resource Management*, 33(6), 1209–1236.

Vrontis, D., Christofi, M., Pereira, V., Tarba, S., Makrides, A., & Trichina, E. (2022). Artificial intelligence, robotics, advanced technologies and human resource management: A systematic review. *The International Journal of Human Resource Management*, 33(6), 1237–1266.

Floridi, L., & Cowls, J. (2019). *A Unified Framework of Five Principles for AI in Society*. Harvard Data Science Review.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). *The Ethics of Algorithms: Mapping the Debate*. Big Data & Society.

Jobin, A., Ienca, M., & Vayena, E. (2019). *The Global Landscape of AI Ethics Guidelines*. Nature Machine Intelligence.

Floridi, L., et al. (2018). *AI4People—An Ethical Framework for a Good AI Society*. Minds and Machines.

European Commission. (2019). *Ethics Guidelines for Trustworthy AI*.

OECD. (2019). *OECD Principles on Artificial Intelligence*.

Goldman, B., & Cropanzano, R. (2015). "Justice" and "fairness" are not the same thing. *Journal of Organizational Behavior*, 36(2), 313–318. <https://doi.org/10.1002/job.1956>

Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationships. In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). Plenum Press. https://doi.org/10.1007/978-1-4613-3087-5_2

Mirowska, A., & Mesnet, L. (2022). Preferring the devil you know: Potential applicant reactions to artificial intelligence evaluation of interviews. *Human Resource Management Journal*, 32(2), 460–479. <https://doi.org/10.1111/1748-8583.12393>

Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2021). Algorithmic management of work on online labor platforms: When matching meets control. *MIS Quarterly*, 45(4), 1999–2022.

<https://doi.org/10.25300/MISQ/2021/15333>

Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Human Resource Management Review*, 1(1), 61–89. [https://doi.org/10.1016/1053-4822\(91\)90011-Z](https://doi.org/10.1016/1053-4822(91)90011-Z)

Paudyal, P., & Wong, B. L. W. (2018). Algorithmic opacity: Making algorithmic processes transparent through abstraction hierarchy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 192–196. <https://doi.org/10.1177/1541931218621046>