E-ISSN: 3108-4192

APSSHS

Academic Publications of Social Sciences and Humanities Studies
2025, Volume 5, Issue 2, Page No: 40-50
Available online at: https://apsshs.com/

Journal of Applied Organizational Systems and Behavior

Designing Contestability and Employee Appeals Into Artificial Intelligence–Mediated Management

Aisha Bello[1]*, Zainab Sule[1], Ibrahim Musa[2], Grace Adeyemi[3], Ahmed Youssef[4]

1. Department of Contestability and AI Appeals, Faculty of Business, University of Ibadan, Ibadan, Nigeria.
2. Department of AI-Mediated Management, Faculty of Management, University of Ilorin, Ilorin, Nigeria.
3. Department of Employee Rights, Faculty of Psychology, Federal University of Agriculture Abeokuta, Abeokuta, Nigeria.
4. Department of Work Systems, Faculty of Psychology, University of Khartoum, Khartoum, Sudan.

Abstract

Artificial intelligence–mediated management increasingly influences how organizations allocate work, evaluate performance, screen personnel, prioritize cases, recommend actions, and structure managerial judgment. Yet governance arrangements often emphasize model accuracy, transparency, or nominal human oversight without specifying how employees can challenge consequential decisions, obtain relevant evidence, reach an authorized reviewer, and secure a reasoned disposition or remedy. This Original Governance Framework Article develops a proposed architecture for organizational contestability in AI-mediated management. It conceptualizes contestability as a practical governance capacity linking notice, challenge grounds, explanation, evidence access, appeal, escalation, consequential human review, remediation, case closure, and organizational learning. The article distinguishes contestability from generic transparency, employee voice, perceived fairness, and symbolic human involvement. It further proposes that meaningful appeal requires more than an opportunity to speak: employees must be able to identify a decision, formulate a challenge, access relevant information, reach an accountable review pathway, and obtain review by an actor with sufficient competence and authority. The framework also separates human presence from remedial authority and explanation from evidentiary sufficiency. Its propositions are intended to organize future empirical inquiry rather than establish causal effects or implementation readiness. The contribution is therefore a testable governance synthesis for analyzing whether AI-mediated managerial decisions are practically challengeable, institutionally reviewable, and capable of producing accountable organizational response under bounded technological, occupational, and institutional conditions.

Keywords: Why contestability matters, Conceptual and governance foundations, Grounds for employee challenge, Explanation and evidence access, Contestability-governance framework, Implementation and audit criteria

How to cite this article: Bello A, Sule Z, Musa I, Adeyemi G, Youssef A. Designing Contestability and Employee Appeals Into Artificial Intelligence–Mediated Management. J Appl Organ Syst Behav. 2025;5(2):40-50. https://doi.org/10.51847/I6A1ccayKk

Received: 18 September 2025; **Revised:** 19 December 2025; **Accepted:** 20 December 2025
Corresponding author: Aisha Bello
E-mail ✉ aisha.bello@ui.edu.ng

Introduction

Artificial intelligence is no longer confined to peripheral decision support. In contemporary organizations, algorithmic systems may participate in directing workers, evaluating performance, allocating opportunities, monitoring behavior, determining priorities, and structuring managerial attention. Research on algorithms at work therefore depicts algorithmic management as a new terrain of organizational control in which computational systems mediate the relationships through which work is directed and contested [1]. Related work-design scholarship shows that algorithmic management can perform functions traditionally associated with managers and thereby alter job demands, resources, discretion, and the organization of work [2]. These developments make governance questions inseparable from the design of work itself.

The organizational consequences of AI nevertheless resist a single technological-determinist interpretation. A multilevel review of organizational AI research identifies heterogeneous implications across worker attitudes, collaboration, capability, interaction, and control, indicating that effects depend on the form of AI, organizational level, and surrounding practices [3]. Accordingly, the relevant question is not whether AI is intrinsically beneficial or harmful. The more specific governance problem is what happens when an employee is affected by an AI-mediated managerial decision and believes that the decision is mistaken, unfair, insufficiently contextualized, improperly delegated, or unsupported by accessible evidence.

Perceived legitimacy cannot be assumed from the presence of computational consistency. Experimental evidence concerning algorithmic management shows that reactions to algorithmic decisions vary with the nature of the managerial task; perceived fairness, trust, and emotion depend partly on whether the task appears appropriate for algorithmic judgment [4]. Thus, technological capability does not itself establish legitimate managerial authority. Similarly, algorithm-based human-resource decision making may create tensions between organizational efficiency and employees' personal integrity, particularly when systems constrain the capacity to understand, contest, or situate consequential judgments within an employee's lived context [5].

This article argues that these problems require contestability, not merely transparency. Across organization and work scholarship, AI-mediated management is treated as a restructuring of work, control, and worker experience rather than as a neutral decision aid [1-3]. Yet transparency alone does not specify what an affected employee may do after learning that an algorithm contributed to a decision. A system may disclose inputs or logic while leaving the decision practically unchallengeable. Likewise, a nominal human reviewer may exist without having sufficient evidence, independence, competence, or authority to alter an outcome.

The article therefore develops an original proposed Contestability-Governance Framework. It defines contestability as the practical organizational capacity of an affected employee to identify and question a consequential AI-mediated managerial decision, obtain sufficient explanation and relevant evidence to formulate a challenge, reach an authorized review process, and receive a reasoned disposition with an available remedy where warranted. This is a non-empirical governance framework rather than a validated intervention. Its purpose is to connect scattered insights on algorithmic control, organizational justice, explanation, employee voice, human-AI delegation, and accountability into a testable architecture while preserving distinctions between conceptual plausibility, empirical association, organizational practice, and demonstrated causal effect.

Conceptual and governance foundations

A governance analysis of contestability must begin from the organizational properties of learning algorithms rather than treating them as passive calculation tools. Learning algorithms can combine opacity, digitized traces, continuous recomputation, and anticipatory quantification in ways that alter how expertise, boundaries, and control are organized [6]. Contestability is therefore not only a user-interface property. It concerns who may question a computationally mediated claim, which knowledge counts in review, what evidence can be introduced, and which organizational actor possesses authority to revise the consequential decision.

This perspective also avoids reducing governance to a choice between automation and human judgment. The automation–augmentation paradox suggests that substitution and complementarity are intertwined organizational processes rather than mutually exclusive technological strategies [7]. Human-AI decision structures can consequently vary in how decision authority is delegated, sequenced, or combined, with their appropriateness depending on characteristics of the decision and organizational context [8]. A governance framework must therefore distinguish technological capability from legitimate authority and distinguish the existence of human participation from the quality and consequences of that participation.

Contestability provides a useful organizing concept because it focuses on responsiveness to human intervention. Existing scholarship conceptualizes contestable AI as systems designed so that relevant actors can intervene across stages of development and use rather than merely observe algorithmic processes [9]. The present article extends that idea into employment governance. It proposes that organizational contestability requires a connected chain of notice, challenge grounds, evidence access, appeal, escalation, consequential review, disposition, remedy, and feedback. That chain is an original synthesis; the literature does not establish that these elements form a validated causal sequence.

The framework is also explicitly multilevel. An employee may experience a decision as opaque or procedurally unfair, but such a perception does not by itself establish organizational-level dysfunction. Conversely, a formally documented appeal policy does not establish that employees can use it safely or consequentially. Evaluation must therefore separate individual perceptions from organizational arrangements, formal policy from enacted practice, technical explanation from usable evidence, and conceptual governance principles from measured organizational effects.

Grounds for employee challenge

Contestability becomes operational only when an employee can identify what is challengeable. Organizational-justice perspectives on employee-managing AI indicate that fairness extends beyond the distributive outcome of a decision to include procedures, interpersonal treatment, informational adequacy, and opportunities for redress [10]. A governance framework

should therefore avoid defining a legitimate challenge as merely an allegation that an algorithm produced an objectively incorrect result. Employees may have reasons to contest how information was gathered, how a rule was applied, whether context was excluded, or whether decision authority was appropriately delegated.

Personnel-selection research reinforces the legitimacy of several such grounds. Reviews of AI-enabled recruitment and selection identify concerns involving bias, fairness, discrimination, privacy, transparency, and accountability [11]. At the same time, systematic evidence on algorithmic fairness perceptions shows substantial heterogeneity across domains, operationalizations, and contexts [12]. Contestability should therefore not presume that employees share one fairness standard or that subjective unfairness automatically proves technical or distributive error. Challenge procedures instead need to make the basis of disagreement explicit enough to permit examination.

Human-resource management creates additional grounds for caution because AI applications depend on data quality, complex organizational objectives, accountability arrangements, and difficult causal assumptions [13]. An apparently precise output may therefore remain contestable when relevant data are incorrect, a model is applied outside an appropriate context, an inference omits consequential information, or decision authority has been transferred without an adequate governance basis.

Building from these literatures, this article proposes eight non-exclusive grounds for employee challenge: data or provenance error; contextual or inference error; distributive or procedural unfairness; discrimination or ethical concern; explanation or evidence insufficiency; inappropriate automation or delegation; inconsistent application; and appeal-process or remedy failure. These categories organize heterogeneous concerns; they do not constitute an empirically validated taxonomy, prevalence hierarchy, or legal classification. Different grounds may overlap, and the relevance of each is conditional on the technology, decision, occupation, institutional setting, and available organizational authority.

Table 1 clarifies the conceptual architecture of contestability by distinguishing the constructs required for an employee challenge to become procedurally intelligible, reviewable, and potentially consequential.

Table 1. Definitions, constructs, and conceptual boundaries for Designing Contestability and Employee Appeals Into Artificial Intelligence–Mediated Management

Construct or component	Definition	Problem addressed	Evidence basis	Distinction from adjacent concepts	Proposed role	Uncertainty	Boundary statement
AI-mediated managerial decision	A consequential managerial judgment, recommendation, allocation, evaluation, or control process materially shaped by an AI or algorithmic system	Employees may experience consequences without knowing how computational mediation entered the decision	Established organizational and algorithmic-management literature	Differs from ordinary digitization or passive information systems because the system materially informs managerial judgment or action	Defines the entry condition for contestability	Degree of algorithmic influence may be difficult to observe in hybrid decisions	Computational involvement does not mean the algorithm independently determined the outcome
Contestability	Practical capacity to question, examine, review, and potentially alter a consequential AI-mediated decision	Transparency can reveal information without creating an actionable challenge pathway	Established concept; workplace architecture is proposed	Broader than explanation but narrower than general participation or employee involvement	Organizing governance principle connecting challenge to consequential response	No validated organizational measure or universally accepted implementation form exists	Contestability does not imply that every challenged decision should be changed
Challenge ground	A specific reason why a decision warrants reconsideration or investigation	Undifferentiated complaints are difficult to route and evaluate	Evidence-informed, article-specific taxonomy	Differs from dissatisfaction because it identifies the alleged basis of decision deficiency	Converts concern into an examinable governance claim	Different grounds may overlap or require different expertise	A challenge ground indicates reviewability, not proof of error
Data or provenance challenge	Claim that relevant data are inaccurate, incomplete, outdated, misattributed, or improperly sourced	Incorrect inputs may propagate through otherwise consistent decision procedures	Evidence-informed category	Concerns informational foundations rather than the model's normative fairness	Enables correction of decision inputs and data lineage	Complete provenance may not always be available to employees	Data correction does not automatically invalidate the full decision
Context or inference challenge	Claim that the system or decision process omitted relevant circumstances or	Standardized models may underrepresent situational information	Proposed governance category grounded in	Differs from raw data error because information may be accurate yet	Creates a route for contextual evidence to enter review	Determining what context is materially relevant may be contested	Contextual disagreement does not itself establish model failure

	drew an inappropriate inference		organizational AI concerns	interpreted inadequately			
Fairness or discrimination challenge	Claim that a decision or process produces unjustified differential treatment or procedurally objectionable consequences	Apparent consistency may coexist with unequal or contested treatment	Established fairness and justice concerns; taxonomy placement proposed	Distinct from simple unfavorable outcome because it questions the legitimacy of treatment or process	Enables scrutiny of distributive, procedural, or discriminatory concerns	Fairness standards may differ across stakeholders, jurisdictions, and decision types	Perceived unfairness and legally established discrimination are not equivalent
Explanation	Intelligible account of the factors, reasoning, or process associated with a decision	Affected employees may be unable to understand why an outcome occurred	Established explanatory construct	Differs from transparency because useful explanation is selective and audience-dependent	Supports comprehension and formulation of a challenge	More information does not necessarily improve understanding	Explanation does not establish evidentiary sufficiency or decision correctness
Evidence access	Access to records, relevant inputs, contextual information, decision rules, provenance, or traceable decision history needed to evaluate a challenge	Appeals may be formally available but impossible to substantiate	Proposed governance extension	Differs from explanation because evidence must permit examination rather than merely comprehension	Provides the evidentiary basis for contestation and review	Access may conflict with privacy, security, confidentiality, or third-party interests	Meaningful access need not imply unrestricted disclosure
Employee appeal	A case-based request for reconsideration of a consequential decision	General voice mechanisms may not create responsibility for response	Proposed formalization informed by employee-voice scholarship	More structured than general speaking up because it creates a recognizable case and expected disposition	Converts challenge into an organizational review obligation	Appropriate formality may vary with decision stakes	Appeal access does not guarantee reversal or remedy
Escalation	Transfer of an unresolved, conflicted, high-stakes, or potentially systemic case to a different or higher review authority	First-line reviewers may lack competence, independence, or remedial power	Proposed governance mechanism	Different from repeated complaint because authority, expertise, or review level changes	Prevents contestability from terminating at a powerless review stage	Triggers and escalation depth require contextual validation	Not every unsuccessful appeal warrants escalation
Consequential human review	Human examination capable of interrogating evidence and exercising legitimate authority over the contested decision	Nominal human presence may produce symbolic rather than substantive oversight	Evidence-informed construct; integrated definition proposed	Differs from "human in the loop" because presence alone is insufficient	Connects contestation to accountable judgment	Reviewer competence, independence, and cognitive biases remain unresolved	Human review does not guarantee superior accuracy or fairness
Remedial authority	Legitimate capacity to affirm, modify, revoke, re-run, correct, or otherwise address a contested decision or process	Review without power to change consequences may be procedurally hollow	Original governance specification	Distinct from advisory review because the reviewer can affect the consequential outcome or responsible process	Makes contestability potentially consequential	Appropriate remedy depends on the source and seriousness of the problem	Availability of a remedy does not imply that reversal is appropriate

Table note. The integrated taxonomy and governance roles are original proposed synthesis. The table organizes conceptual boundaries and does not establish legal entitlements, validated causal relationships, prevalence rankings, or universal design requirements.

Explanation and evidence access

Explanation is necessary to contestability but insufficient on its own. Field research on professionals using opaque AI for consequential diagnostic judgments shows that opacity can alter how experts engage with algorithmic recommendations when

they remain responsible for critical decisions [14]. The relevance to employee appeals is conceptual rather than directly empirical: an employee or reviewer cannot meaningfully test a contested decision merely because a system provides a label or recommendation. What matters is whether the information made available permits the decision to be situated, interrogated, and compared with relevant alternative evidence.

Evidence from personnel-selection contexts likewise indicates that explanation is not a context-free governance solution. People's reactions to algorithmic versus human decision makers can vary with whether explanations are provided and with the kind of selection test involved [15]. Separate experimental evidence indicates that assuring individuals that humans remain involved in AI-supported decisions can influence trust [16]. These findings support treating explanation and human involvement as consequential perceptions, but neither establishes that an explanation is evidentially sufficient or that nominal human oversight produces accurate, fair, or remedial review.

Explainability research further suggests that perceptions of explainability and causability are associated with trust and acceptance [17]. Yet trust is not the same as legitimacy, and acceptance is not evidence that a decision is correct. A contestability architecture must therefore distinguish explanation from evidence access. Explanation concerns intelligibility: what factors or reasoning generated a decision. Evidence access concerns what an employee or reviewer can inspect to evaluate the challenge—for example, relevant input records, provenance, applicable rules, contextual information, or a preserved decision trail.

This distinction also creates boundaries. Full disclosure may be impossible where information implicates privacy, security, confidential business information, or third parties. Conversely, withholding all relevant material can make an appeal procedurally hollow. The proposed framework therefore treats evidentiary sufficiency as contextual rather than absolute: an affected employee must receive enough information to formulate and support a specific challenge, while access remains bounded by legitimate competing obligations. Whether particular forms or levels of access improve error correction, perceived fairness, trust calibration, or organizational learning remains an empirical question.

Appeals and escalation pathways

Contestability requires a pathway that converts speaking up into an organizationally recognized case. Employee voice is shaped by perceived risk, motives, relationships, and context rather than formal permission alone [18]. Meta-analytic evidence also distinguishes promotive from prohibitive voice, showing different antecedents and associations [19]. An appeal should therefore not be treated as generic voice. It is a structured challenge to a consequential decision that should generate case ownership, review, and a reasoned disposition.

Managerial response is equally consequential. Voice endorsement varies with directness, credibility, and politeness [20]. Digital voice channels likewise enable or constrain speaking up through channel affordances and managerial reactions [21]. These findings support a distinction between channel availability and consequential access. The proposed framework therefore makes appeal usability dependent on identifiable routing, case status, response obligations, and escalation where first-line review is insufficient.

A further distinction concerns escalation. Escalation is not simply repeated voice or an automatic second opinion. In the proposed architecture it is a governance response to conditions such as unresolved evidentiary conflict, insufficient first-line authority, potential reviewer conflict, high decision stakes, or indications that a case may reflect a recurring system problem. Whether those conditions warrant separate reviewers, different expertise, or organizational-level investigation is context dependent and requires empirical testing.

Table 2 specifies the mechanisms through which contestability may become substantively consequential while identifying rival explanations and the evidence required to test each proposed relationship.

Table 2. Proposed mechanisms, relationships, and organizational consequences in Designing Contestability and Employee Appeals Into Artificial Intelligence–Mediated Management

Mechanism or relationship	Antecedent	Mediator or process	Moderator	Expected consequence	Supporting evidence	Rival explanation	Validation requirement
Decision opacity → challenge salience	Employee encounters an unexpected or poorly understood AI-mediated outcome	Uncertainty prompts scrutiny of the decision basis	Decision stakes; prior trust; employee expertise	Greater likelihood of seeking clarification or challenge	Evidence base indicates that opacity matters in consequential judgment	Dissatisfaction with the outcome may drive challenge independently of opacity	Compare reactions to equivalent favorable and unfavorable decisions
Perceived unfairness → appeal initiation	Employee interprets outcome or procedure as unjustified	Fairness appraisal converts dissatisfaction	Decision type; organizational justice climate; power asymmetry	Increased probability of formal challenge	Fairness and justice concerns are established in algorithmic	Employees may challenge primarily because the decision is unfavorable	Distinguish fairness reasoning from outcome preference

		into a procedural concern			decision contexts		
Explanation → decision intelligibility	Employee receives information about decision logic or influential factors	Greater comprehension of how the outcome was produced	Complexity; explanation form; employee knowledge	Better ability to identify whether a specific challenge exists	Explanation can alter reactions and understanding	Explanation may merely increase perceived legitimacy without increasing knowledge	Test comprehension separately from satisfaction and trust
Evidence access → challenge quality	Employee receives relevant records or contextual decision information	Ability to compare organizational evidence with alternative evidence	Data sensitivity; information volume; employee capability	More specific and reviewable appeal claims	Opacity literature supports the importance of inspectable information	Information volume may produce overload rather than better contestation	Assess factual specificity and evidentiary relevance of appeals
Accessible channel → appeal usability	Formal appeal mechanism exists	Low procedural burden and understandable routing facilitate entry	Hierarchy; retaliation risk; digital literacy; organizational climate	Greater practical accessibility of contestation	Voice-channel research supports the importance of enacted affordances	High usage may indicate poor underlying decision quality rather than good governance	Measure accessibility separately from appeal frequency
Managerial response → continued participation	Employee submits a challenge	Response quality signals whether speaking up has consequences	Manager credibility; power relations; issue sensitivity	Continued voice, withdrawal, or silence	Voice research shows that managerial response conditions employee reactions	Broader trust in management may determine both response perceptions and future participation	Longitudinally examine response and subsequent challenge behavior
Reviewer competence → review quality	Appeal reaches human reviewer	Reviewer identifies relevant evidence, uncertainty, and system limitations	Domain expertise; AI knowledge; workload	More discriminating review of contested decisions	Human-AI delegation research supports the role of comparative expertise	Expertise may increase overconfidence or algorithm aversion	Measure accuracy, calibration, and reasoning quality separately
Reviewer authority → substantive contestability	Reviewer identifies a credible problem	Formal ability to change outcome or process converts judgment into action	Organizational hierarchy; decision stakes; institutional rules	Review can produce meaningful correction where warranted	Evidence supports distinguishing human involvement from control	Reversal may reflect managerial discretion rather than superior review	Trace authority, reasoning, and disposition independently
Escalation → conflict-sensitive review	First-line review is unresolved, conflicted, or insufficient	Case moves to different expertise or authority	Stakes; systemic implications; reviewer conflict	Greater possibility of resolving cases beyond first-line limitations	Mechanism is evidence-informed but proposed	Escalation may simply add delay and administrative burden	Compare escalated and non-escalated cases while accounting for case complexity
Remedy → organizational learning	Challenge reveals error, recurring pattern, or governance weakness	Case information feeds system, policy, or process review	Record quality; governance ownership; learning capacity	Reduced recurrence or improved decision processes, proposed	Organizational learning logic supports feedback conceptually	Changes may arise from unrelated organizational reforms	Longitudinally link case patterns to documented governance changes
Procedural burden → chilling effect	Appeals require excessive time, documentation, or personal exposure	Expected cost suppresses challenge behavior	Employment insecurity; status; managerial climate	Underuse despite nominal accessibility	Voice and power research make this mechanism plausible	Low appeal levels may reflect high decision quality	Combine utilization data with employee-reported barriers
Repeated contestability → strategic adaptation	Employees and managers learn how appeal mechanisms operate	Behavioral adaptation changes challenge and decision practices	Incentives; transparency; sanctions	Learning, gaming, defensive decision making, or improved process discipline	Proposed mechanism	Observed behavioral change may reflect unrelated policy changes	Use longitudinal and comparative designs

Table note. "Expected consequence" identifies a proposed relationship or evaluation target, not an established effect of the framework. Rival explanations are included to prevent causal interpretation from cross-sectional patterns.

Human review and remedial authority

Human review is meaningful only when it contributes more than symbolic presence. Human-AI symbiosis scholarship identifies potential complementarity between computational analysis and contextual judgment [22]. Experimental evidence shows that productive delegation requires metaknowledge about comparative human and AI strengths and errors [23]. Workplace evidence further indicates that domain experience can simultaneously improve ability and shape algorithm aversion or accountability concerns [24].

Reviewer competence is therefore a governance question rather than a job-title assumption. Reviewers may need domain knowledge, access to the decision record, and understanding of system limitations. Yet competence without authority can leave contestability hollow. Experiments show that even limited ability to modify algorithmic output can change willingness to use it [25]. This does not validate employee override rights, but it distinguishes consequential control from observation.

This distinction matters because review quality has at least two dimensions: epistemic capacity and consequential authority. A reviewer may understand the model but lack authority to alter the employment decision, or possess formal authority while lacking access to the evidence needed to interrogate the system. The framework treats either configuration as potentially incomplete contestability. It also leaves open whether reviewer independence should increase with stakes, conflict, or systemic significance rather than operating as a uniform requirement across all cases.

The proposed framework defines consequential human review as a bounded combination of competence, evidence access, responsibility, appropriate independence, and remedial authority. Available remedies may include affirming, modifying, revoking, re-running, correcting data, or correcting a process. Human review does not guarantee unbiased or correct decisions; its quality and authority require testing.

Proposed contestability-governance framework

Algorithmic systems can alter which forms of knowledge become authoritative. Qualitative research shows that algorithm introduction may reshape a regime of knowing [26]. Platform research likewise links algorithmic matching with managerial control and worker tensions [27], while employment-relations scholarship describes algorithmic management as reorganizing work allocation and performance processes [28]. Mixed-method gig-work evidence further shows that experienced autonomy can coexist with algorithmic control [29]. Formal discretion therefore cannot be assumed to provide substantive power to challenge decisions.

The proposed process links an identifiable consequential decision with a contestability trigger, notice, explanation and evidence access, challenge-ground specification, first-line appeal, escalation where warranted, consequential human review, disposition or remedy, case closure, and feedback into governance learning. Accessibility, non-retaliation, traceability, reviewer competence, proportional independence, and timeliness are proposed cross-cutting conditions.

Table 3 translates the proposed framework into falsifiable governance propositions and evaluation criteria while separating organizational responsibility from evidence of actual readiness.

Table 3. Testable propositions, evaluation criteria, and practical responsibilities

Proposition or criterion	Unit and context	Observable indicator	Evidence required	Falsification condition	Organizational responsibility	Misuse risk	Readiness boundary
P1. Decision noticeability	Employee facing a consequential AI-mediated decision	Employee can identify that algorithmic mediation materially informed the decision	Decision notices; process documentation; employee comprehension evidence	Employees cannot reliably determine whether AI materially shaped the decision	Decision owner	Technical disclosure presented as meaningful notice	Notice is necessary but insufficient for contestability
P2. Ground specificity	Individual appeal case	Employee can state a recognizable basis for challenge	Appeal content; routing records; reviewer classification	Cases cannot be translated into examinable claims	Appeal-process owner	Taxonomy used to exclude unconventional but legitimate concerns	Categories must remain revisable
P3. Evidentiary sufficiency	Appeal and review episode	Available information permits a specific claim to be investigated	Explanations; data lineage; decision records; contextual evidence	Review remains impossible despite nominal disclosure	System and data owners	Information dumping or selective disclosure	No universal evidence package can be assumed
P4. Accessible initiation	Employee–organization interface	Eligible employee can initiate a case without unreasonable procedural burden	Process mapping; employee experience; abandonment data	Material barriers prevent practical access	Governance owner	Low appeal volume interpreted automatically as success	Accessibility must be assessed independently from use frequency

P5. Non-retaliatory use	Employee and managerial context	Challenge does not predict punitive or suppressive treatment attributable to appeal behavior	Longitudinal employment and process evidence	Credible challenge is followed by systematic retaliatory consequences	Senior governance and HR functions	Formal anti-retaliation statements treated as proof	Requires behavioral rather than policy-only evidence
P6. Reviewer competence	Human review episode	Reviewer can identify relevant decision assumptions, evidence limitations, and contextual alternatives	Reviewer reasoning; expertise records; calibrated judgment tests	Reviewers cannot distinguish material from irrelevant challenge information	Review-function owner	Credentials substituted for demonstrated competence	No fixed universal credential threshold
P7. Review independence proportional to stakes	High-stakes or conflicted appeal	Reviewer is sufficiently separated from interests implicated in the initial decision	Governance structure; conflict records; escalation decisions	Same conflicted actor controls both initial decision and final review without safeguard	Governance authority	Independence treated as complete organizational detachment	Required degree depends on risk and context
P8. Consequential remedial authority	Review outcome	Authorized reviewer can implement an appropriate disposition	Authority matrix; disposition record; remedy execution	Reviewer can identify error but cannot affect outcome or responsible process	Decision authority	Symbolic "human oversight" without consequential power	Authority should match decision significance
P9. Escalation adequacy	Unresolved, systemic, or high-stakes case	Case reaches appropriately differentiated authority or expertise	Routing history; escalation rationale; review record	Credible unresolved concerns terminate solely because first-line authority is exhausted	Governance owner	Escalation becomes procedural delay	Escalation is conditional rather than mandatory
P10. Traceable case closure	Completed appeal	Final disposition, reasoning, actions, and unresolved uncertainty are recorded	Case file; decision trail; communication record	Organization cannot reconstruct how the challenge was resolved	Appeal-process owner	Administrative closure mistaken for substantive resolution	Documentation quality requires qualitative examination
P11. Governance learning	Organizational or system level	Recurrent challenge patterns trigger review of system, policy, data, or decision practice	Aggregated case records; documented corrective actions	Recurrent materially similar failures produce no organizational examination	AI governance function	Appeal counts converted into simplistic performance scores	Requires longitudinal evidence linking patterns to response
P12. Adverse-effect monitoring	Organizational implementation	Burden, delay, chilling, strategic gaming, and overreliance are examined alongside intended outcomes	Longitudinal employee, process, and organizational evidence	Governance assessment records only favorable indicators	Independent audit or governance function	Framework becomes self-confirming	Implementation cannot be considered mature without counterevidence testing

Table note. All propositions and readiness boundaries are original proposed synthesis. Observable indicators specify what could be investigated, not validated metrics or minimum standards. No proposition establishes legal compliance, organizational effectiveness, certification, causal benefit, or deployment readiness.

Implementation and audit criteria

Contestability can fail if organizations evaluate policy existence rather than enacted practice. Ethics-based auditing has been conceptualized as a structured governance process constrained by technical, social, economic, organizational, and institutional conditions [30]. Evaluations of AI ethics guidelines likewise show that principles do not automatically specify implementation practices [31]. Auditing contestability should therefore examine whether employees can actually progress from notice to review and disposition.

Figure 1 presents the original conceptual synthesis for mechanisms, boundary conditions, and organizational consequences, showing how the article's principal constructs, mechanisms, uncertainties, and decision boundaries are connected.

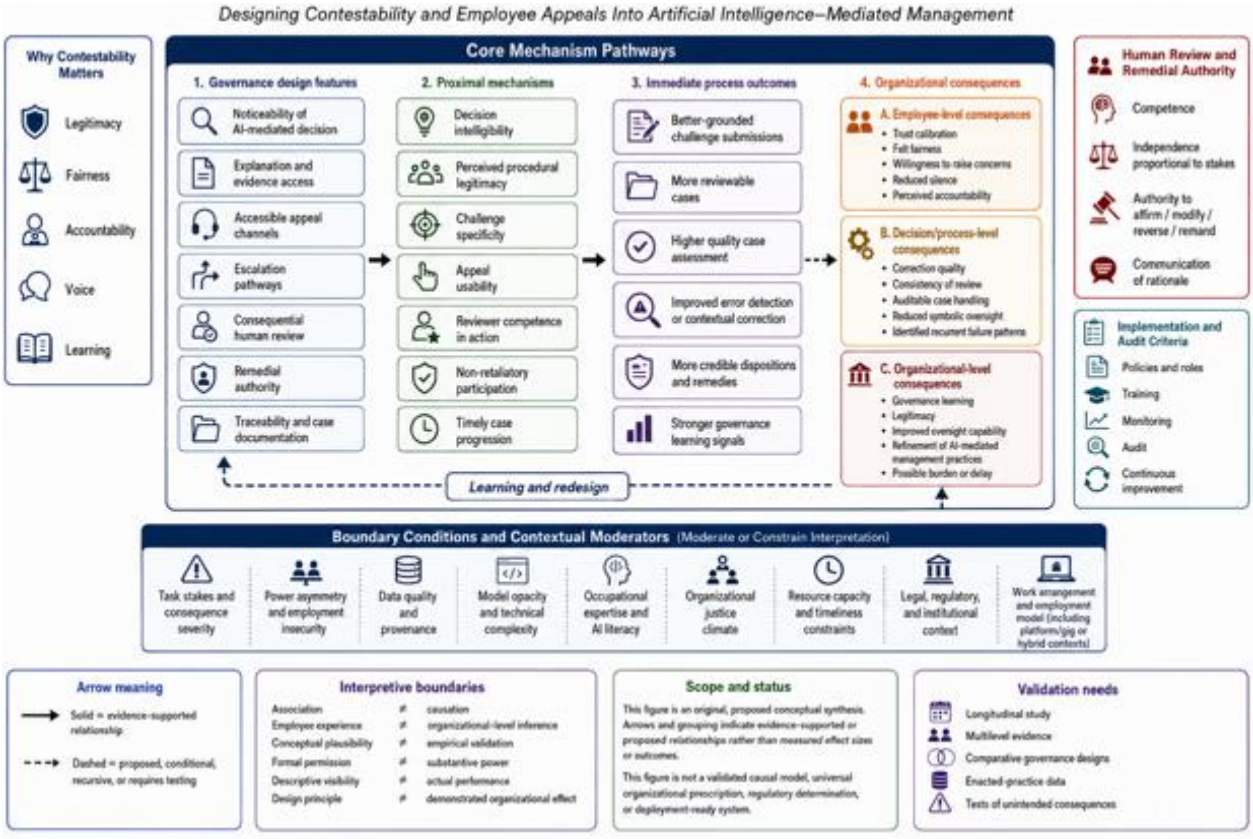

Figure 1. Mechanisms, Boundary Conditions, and Organizational Consequences

Accountability cannot be displaced onto an algorithm. Organizational actors remain implicated in the value-laden choices embedded in algorithmic systems [32], while practice-oriented work shows that translating principles into action requires concrete tools and methods [33]. Operational AI accountability therefore requires more than statements of principle because governance also depends on auditable processes, practical implementation, and identifiable responsibility [30-32].

Proposed audit domains include noticeability, evidentiary sufficiency, accessibility, timeliness, reviewer competence, appropriate independence, remedial authority, traceability, consistency, closure, recurrence, and learning. Evaluation should also test delay, burden, chilling effects, strategic gaming, and symbolic review. These are proposed dimensions, not certification thresholds or proof of readiness.

Conclusion

This article proposes contestability as an organizational governance capacity rather than a synonym for transparency, explanation, employee voice, or human oversight. The Contestability-Governance Framework links identifiable AI-mediated decisions with challenge grounds, explanation and evidence access, appeal, escalation, consequential human review, remedial authority, closure, and governance feedback.

The framework separates opportunities to speak from consequential influence, explanation from evidence sufficiency, human presence from review authority, technological capability from legitimate authority, and formal policy from enacted practice. Its categories, pathways, propositions, and audit criteria have not been validated as an integrated system. Their relevance is likely to vary with task characteristics, decision stakes, expertise, organizational power, employment arrangements, institutional settings, and technology.

The contribution is therefore a testable governance synthesis: organizations and researchers can ask not merely whether an AI-mediated managerial system is transparent or supervised, but whether affected employees can challenge it in ways that are intelligible, evidentially grounded, reviewable, consequential, and auditable.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Kellogg KC, Valentine MA, Christin A. Algorithms at work: The new contested terrain of control. Acad Manag Ann. 2020;14(1):366-410. doi:10.5465/annals.2018.0174
2. Parent-Rocheleau X, Parker SK. Algorithms as work designers: How algorithmic management influences the design of jobs. Hum Resour Manag Rev. 2022;32(3):100838. doi:10.1016/j.hrmr.2021.100838
3. Bankins S, Ocampo AC, Marrone M, Restubog SLD, Woo SE. A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. J Organ Behav. 2024;45(2):159-82. doi:10.1002/job.2735
4. Lee MK. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data Soc. 2018;5(1):2053951718756684. doi:10.1177/2053951718756684
5. Leicht-Deobald U, Busch T, Schank C, Weibel A, Schafheitle S, Wildhaber I, et al. The challenges of algorithm-based HR decision-making for personal integrity. J Bus Ethics. 2019;160(2):377-92. doi:10.1007/s10551-019-04204-w
6. Faraj S, Pachidi S, Sayegh K. Working and organizing in the age of the learning algorithm. Inf Organ. 2018;28(1):62-70. doi:10.1016/j.infoandorg.2018.02.005
7. Raisch S, Krakowski S. Artificial intelligence and management: The automation–augmentation paradox. Acad Manag Rev. 2021;46(1):192-210. doi:10.5465/amr.2018.0072
8. Shrestha YR, Ben-Menahem SM, von Krogh G. Organizational decision-making structures in the age of artificial intelligence. Calif Manage Rev. 2019;61(4):66-83. doi:10.1177/0008125619862257
9. Alfrink K, Keller I, Kortuem G, Doorn N. Contestable AI by design: Towards a framework. Minds Mach. 2023;33(4):613-39. doi:10.1007/s11023-022-09611-z
10. Robert LP, Pierce C, Marquis L, Kim S, Alahmad R. Designing fair AI for managing employees in organizations: A review, critique, and design agenda. Hum Comput Interact. 2020;35(5-6):545-75. doi:10.1080/07370024.2020.1735391
11. Hunkenschroer AL, Luetge C. Ethics of AI-enabled recruiting and selection: A review and research agenda. J Bus Ethics. 2022;178(3):977-1007. doi:10.1007/s10551-022-05049-6
12. Starke C, Baleis J, Keller B, Marcinkowski F. Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. Big Data Soc. 2022;9(2):20539517221115189. doi:10.1177/20539517221115189
13. Tambe P, Cappelli P, Yakubovich V. Artificial intelligence in human resources management: Challenges and a path forward. Calif Manage Rev. 2019;61(4):15-42. doi:10.1177/0008125619867910
14. Lebovitz S, Lifshitz-Assaf H, Levina N. To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. Organ Sci. 2022;33(1):126-48. doi:10.1287/orsc.2021.1549
15. Wesche JS, Hennig F, Kollhed CS, Quade J, Kluge S, Sonderegger A. People's reactions to decisions by human vs. algorithmic decision-makers: The role of explanations and type of selection tests. Eur J Work Organ Psychol. 2024;33(2):146-57. doi:10.1080/1359432X.2022.2132940
16. Aoki N. The importance of the assurance that "humans are still in the decision loop" for public trust in artificial intelligence: Evidence from an online experiment. Comput Human Behav. 2021;114:106572. doi:10.1016/j.chb.2020.106572
17. Shin D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. Int J Hum Comput Stud. 2021;146:102551. doi:10.1016/j.ijhcs.2020.102551
18. Morrison EW. Employee voice and silence: Taking stock a decade later. Annu Rev Organ Psychol Organ Behav. 2023;10:79-107. doi:10.1146/annurev-orgpsych-120920-054654
19. Chamberlin M, Newton DW, LePine JA. A meta-analysis of voice and its promotive and prohibitive forms: Identification of key associations, distinctions, and future research directions. Pers Psychol. 2017;70(1):11-71. doi:10.1111/peps.12185
20. Lam CF, Lee C, Sui Y. Say it as it is: Consequences of voice directness, voice politeness, and voicer credibility on voice endorsement. J Appl Psychol. 2019;104(5):642-58. doi:10.1037/apl0000358
21. Ellmer M, Reichel A. Mind the channel! An affordance perspective on how digital voice channels encourage or discourage employee voice. Hum Resour Manag J. 2021;31(1):259-76. doi:10.1111/1748-8583.12297
22. Jarrahi MH. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Bus Horiz. 2018;61(4):577-86. doi:10.1016/j.bushor.2018.03.007
23. Fügener A, Grahl J, Gupta A, Ketter W. Cognitive challenges in human–artificial intelligence collaboration: Investigating the path toward productive delegation. Inf Syst Res. 2022;33(2):678-96. doi:10.1287/isre.2021.1079

24. Allen RT, Choudhury P. Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. Organ Sci. 2022;33(1):149-69. doi:10.1287/orsc.2021.1554
25. Dietvorst BJ, Simmons JP, Massey C. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Manage Sci. 2018;64(3):1155-70. doi:10.1287/mnsc.2016.2643
26. Pachidi S, Berends H, Faraj S, Huysman M. Make way for the algorithms: Symbolic actions and change in a regime of knowing. Organ Sci. 2021;32(1):18-41. doi:10.1287/orsc.2020.1377
27. Möhlmann M, Zalmanson L, Henfridsson O, Gregory RW. Algorithmic management of work on online labor platforms: When matching meets control. MIS Q. 2021;45(4):1999-2022. doi:10.25300/MISQ/2021/15333
28. Duggan J, Sherman U, Carbery R, McDonnell A. Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. Hum Resour Manag J. 2020;30(1):114-32. doi:10.1111/1748-8583.12258
29. Wood AJ, Graham M, Lehdonvirta V, Hjorth I. Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. Work Employ Soc. 2019;33(1):56-75. doi:10.1177/0950017018785616
30. Mökander J, Morley J, Taddeo M, Floridi L. Ethics-based auditing of automated decision-making systems: Nature, scope, and limitations. Sci Eng Ethics. 2021;27(4):44. doi:10.1007/s11948-021-00319-4
31. Hagendorff T. The ethics of AI ethics: An evaluation of guidelines. Minds Mach. 2020;30(1):99-120. doi:10.1007/s11023-020-09517-8
32. Martin K. Ethical implications and accountability of algorithms. J Bus Ethics. 2019;160(4):835-50. doi:10.1007/s10551-018-3921-3
33. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Sci Eng Ethics. 2020;26(4):2141-68. doi:10.1007/s11948-019-00165-5