Towards a Theory of Hybrid Accountability: Rethinking Performance Measurement and Organizational Governance in AI-Augmented Organizations

The increasing integration of Artificial Intelligence (AI) into organizational decision-making is reshaping performance measurement, managerial control, and responsibility. While AI increasingly influences organizational decisions, existing Performance Measurement Systems (PMS) and accountability frameworks remain largely centred on identifiable human decision-makers. This creates an Organizational Accountability Gap, whereby decision influence becomes distributed across human and algorithmic contributors while formal accountability remains human and organizational. Using an integrative conceptual review and theory-building approach, this study synthesizes research on performance measurement, organizational accountability, organizational governance, and AI-enabled decision-making. It develops the Theory of Hybrid Accountability, conceptualizing accountability as an organizational governance capability enacted through four mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. These mechanisms are integrated within a Performance Accountability Architecture (PAA) that embeds accountability throughout hybrid decision processes. The framework advances accountability theory, extends performance measurement through a governance-oriented perspective, and broadens AI governance from the governance of algorithms to the governance of AI-enabled organizations. It provides a theoretical foundation for preserving meaningful human and organizational accountability under conditions of distributed decision influence.

Keywords: *Hybrid Accountability; Artificial Intelligence; Performance Measurement Systems; Organizational Governance; AI-Augmented Organizations; Organizational Accountability Gap; Management Control; Performance Accountability Architecture.*

ONLY FOR REVIEW

Introduction

Artificial Intelligence (AI) is reshaping organizational decision-making by generating predictions, recommendations, classifications, and probabilistic assessments that increasingly influence strategic and operational choices (Jarrahi, 2018; Rai, 2020; Raisch & Krakowski, 2021). AI systems now contribute to forecasting, budgeting, performance evaluation, resource allocation, risk management, and strategic planning, signalling a transition toward "AI-augmented organizations", where outcomes increasingly emerge from interactions between human expertise and algorithmic contributions. As organizations rely on AI to interpret complex environments and recommend actions, decision influence becomes distributed across socio-technical systems in which human judgment and algorithmic contributions jointly shape outcomes (Faraj, Pachidi, & Sayegh, 2018; Dellermann et al., 2019; Raisch & Krakowski, 2021). This transformation challenges established assumptions underlying organizational governance and management control. Accountability is particularly affected by this transformation. It has traditionally been understood as the obligation of organizational actors to explain, justify, and assume responsibility for their decisions and consequences (Roberts, 1991; Sinclair, 1995; Bovens, 2007). Within management accounting and performance measurement, accountability is closely associated with systems of planning,

measurement, evaluation, and control that attribute organizational outcomes to identifiable decision-makers (Kaplan & Norton, 1996; Ferreira & Otley, 2009; Franco-Santos et al., 2012). These perspectives implicitly assume an alignment among decision authority, responsibility, and performance evaluation. AI-augmented decision-making increasingly challenges this alignment. Although managers formally retain decision authority, AI systems can materially influence organizational choices by generating predictions, prioritizing alternatives, evaluating risks, and recommending actions. Managers may therefore remain accountable for decisions substantially shaped by algorithmic contributions, while AI systems exercise decision influence without becoming accountable organizational actors. We conceptualize this condition as the Organizational Accountability Gap: a structural misalignment between the distribution of decision influence across human and algorithmic contributors and the formal allocation of accountability for organizational outcomes. The gap is not primarily a question of whether AI itself should be held responsible; it is an organizational governance problem arising when accountability remains human and organizational while decision processes become increasingly hybrid. This problem lies at the intersection of several research streams. AI governance research emphasizes explainability, transparency, fairness, accountability, auditing, and regulatory compliance as conditions for trustworthy and responsible AI (Floridi & Cowls, 2019; Novelli, Taddeo, & Floridi, 2024). More recently, management scholarship has highlighted the need to align control and accountability among stakeholders involved in AI development, deployment, and organizational use. Performance measurement and management control research, meanwhile, explains how organizations design measurement systems to align managerial behaviour with strategic objectives, support organizational learning, and improve performance (Ferreira & Otley, 2009; Franco-Santos et al., 2012). Yet performance measurement remains predominantly centred on human decision-makers, while AI governance focuses primarily on the governance of technological systems. What remains underdeveloped is an integrative explanation of how accountability should be organized when consequential decision influence is distributed across human and algorithmic contributors while formal authority and accountability remain human and organizational. Accordingly, this study addresses the following research question: How should organizational accountability be reconceptualized when decision influence is distributed across human and algorithmic contributors while formal decision authority and accountability remain human and organizational? To answer this question, the study adopts an integrative conceptual review and theory-building approach. It synthesizes four complementary bodies of research—performance measurement systems, organizational accountability, organizational governance, and AI-enabled decision-making—to identify their shared assumptions concerning decision authority, responsibility, performance evaluation, and organizational control. Building on this synthesis, the paper develops the Theory of Hybrid Accountability, which conceptualizes accountability as an organizational governance capability through which algorithmic transparency, human oversight, decision authority, and responsibility allocation are coordinated to preserve meaningful accountability across hybrid decision processes characterized by distributed decision influence. AI systems are therefore not conceptualized as accountable organizational actors; rather, the theory explains how human and organizational answerability can remain meaningful when algorithmic contributions materially influence decisions. The paper makes three principal theoretical contributions. First, it extends accountability theory by conceptualizing accountability as an organizational governance capability capable of preserving human and organizational

ONLY FOR REVIEW

answerability under conditions of distributed decision influence. Second, it bridges performance measurement and AI governance by explaining how consequential algorithmic influence reshapes the relationship among decision authority, responsibility, and organizational control. Third, it introduces the Performance Accountability Architecture (PAA) as a framework for embedding accountability within performance measurement and control in AI-augmented organizations. The remainder of the paper develops the theoretical background and emerging theoretical puzzle, presents the Theory of Hybrid Accountability and the PAA together with their governance mechanisms and theoretical propositions, and discusses the framework's theoretical and managerial implications, limitations, and directions for future research.

Theoretical Background And The Emerging Theoretical Puzzle

Performance Measurement as a Human-Centred Control System

Performance Measurement Systems (PMS) constitute a central component of organizational governance, enabling organizations to monitor performance, evaluate managerial actions, allocate resources, and align behaviour with strategic objectives (Neely, Gregory, & Platts, 1995; Kaplan & Norton, 1996; Ferreira & Otley, 2009). Although performance measurement has evolved from predominantly financial indicators toward multidimensional frameworks incorporating operational, strategic, behavioural, and sustainability dimensions (Franco-Santos, Lucianetti, & Bourne, 2012), its underlying governance logic remains largely human-centred. Across management accounting and performance management research, PMS support managerial judgment, strategic alignment, organizational learning, and control while connecting organizational outcomes to identifiable human actors (Simons, 1995; Ferreira & Otley, 2009; Merchant & Van der Stede, 2017). Performance indicators therefore both inform managerial decisions and provide a basis for evaluating and holding managers accountable. This logic assumes that decision influence, managerial action, performance outcomes, and accountability can ultimately be connected to identifiable human decision-makers. The Balanced Scorecard similarly links strategic objectives and organizational outcomes to managerial actions (Kaplan & Norton, 1996), while Ferreira and Otley's (2009) Performance Management Framework integrates planning, measurement, evaluation, rewards, and feedback. More broadly, management control systems enable managers to exercise judgment, coordinate activities, and implement strategy (Simons, 1995; Merchant & Van der Stede, 2017). Digital transformation expanded the analytical capabilities of PMS without fundamentally altering this governance logic. Information systems, business intelligence, predictive analytics, and real-time dashboards enhanced managerial decision-making while managers continued to exercise judgment and retain organizational responsibility. AI represents a more consequential development because contemporary systems increasingly perform cognitive activities associated with managerial judgment, including forecasting, prioritizing alternatives, and recommending actions (Jarrahi, 2018; Rai, 2020; Raisch & Krakowski, 2021). Algorithmic systems can therefore materially shape how organizational decisions are formulated and evaluated rather than merely provide information to human decision-makers. This development challenges the conventional correspondence among managerial authority, performance, and accountability.

UNDER PEER REVIEW

Organizational outcomes increasingly emerge from interactions between human expertise and algorithmic contributions, while PMS continue largely to operate through governance architectures centred on identifiable human decision-makers. The challenge posed by AI therefore concerns not only the technological capabilities incorporated into PMS, but the governance logic through which organizational performance is evaluated and responsibility attributed. Understanding accountability in AI-augmented organizations consequently requires reconsidering the human-centred assumptions underlying conventional performance measurement systems.

Accountability as the Foundation of Organizational Governance

Accountability is a fundamental principle of organizational governance, ensuring that authority is exercised responsibly, decisions can be justified, and organizational actors remain answerable for their consequences (Roberts, 1991; Sinclair, 1995; Bovens, 2007). Beyond a normative principle, accountability connects decision-making, organizational control, legitimacy, and performance evaluation. Roberts (1991) conceptualizes accountability as a relational process through which actors are called upon to justify their conduct, while Sinclair (1995) emphasizes its multidimensional managerial, political, professional, legal, and personal forms. Bovens (2007) defines accountability as a relationship in which an actor must explain and justify conduct before a forum capable of questioning, evaluating, and potentially imposing consequences. Despite their differences, these perspectives share three assumptions: accountability presupposes an identifiable accountable actor; responsibility can be connected to organizational authority; and actions and outcomes must be sufficiently traceable to permit evaluation. Accountability therefore depends upon the relationship among decision authority, responsibility, traceability, and answerability. Management accounting, internal control, auditing, and performance evaluation similarly rely on attributing organizational actions and outcomes to identifiable actors (Merchant & Van der Stede, 2017). More recent scholarship recognizes that accountability can operate through networks, partnerships, ecosystems, and collaborative arrangements involving multiple actors and overlapping responsibilities (Messner, 2009; Roberts, 2009). Yet this expansion toward relational and collective forms changes the distribution of accountability more readily than the conception of the accountable subject. Whether individual or collective, existing frameworks ultimately locate answerability within human actors, groups, or organizations. Conventional accountability theory therefore remains fundamentally human-centred. AI complicates this logic because algorithmic systems can exert consequential influence within organizational decision processes by generating predictions, classifications, rankings, and recommendations. Organizational decisions may consequently emerge from interactions between managerial judgment and algorithmic outputs. The theoretical problem is not that AI should become an accountable organizational actor, but that it may acquire substantial decision influence without corresponding answerability. Formal responsibility remains assigned to managers and organizations even when decision processes increasingly incorporate algorithmic contributions. This separation exposes a boundary condition of conventional accountability theory. Existing scholarship accommodates the distribution of responsibility among human actors but provides more limited conceptual resources for decision processes that become socio-technical. The challenge posed by AI therefore does not invalidate established accountability theory; rather, it reveals the limits of its explanatory

UNDER PEER REVIEW

reach when decision influence becomes distributed across human and algorithmic contributors while accountability remains institutionally concentrated in human and organizational actors. Understanding governance under these conditions requires extending accountability theory to explain how organizations preserve answerability, responsibility, and control within hybrid decision processes.

AI-Augmented Decision-Making and Distributed Decision Influence

The rapid diffusion of Artificial Intelligence is altering the architecture of organizational decision-making. Unlike earlier information technologies, contemporary AI systems increasingly contribute to cognitive activities associated with managerial judgment by generating forecasts, identifying opportunities, detecting anomalies, and recommending actions (Jarrahi, 2018; Rai, 2020). Their significance lies not in becoming organizational actors, but in exerting consequential influence over how decisions are formulated and evaluated. Research on human–AI collaboration emphasizes complementarity rather than substitution between human expertise and algorithmic capabilities (Dellermann et al., 2019; Raisch & Krakowski, 2021). Human decision-makers contribute contextual understanding, ethical reasoning, creativity, and tacit knowledge, while AI provides computational capacity, pattern recognition, and predictive capabilities. From a socio-technical perspective, organizational decisions therefore emerge through interactions among human actors, technologies, routines, and data infrastructures (Faraj, Pachidi, & Sayegh, 2018). This development can be understood through the concept of distributed decision influence. Algorithmic systems can structure information, frame alternatives, prioritize options, and direct managerial attention (Rai, 2020; Grote, Parker, & Crowston, 2026). Such influence should not be equated with human agency, intentionality, moral responsibility, or formal organizational authority. Rather, decision influence may become distributed across human and algorithmic contributors even when formal decision authority and accountability remain assigned to human actors. This does not imply the displacement of managerial judgment. Effective AI use continues to depend upon meaningful human oversight, particularly under conditions of uncertainty, ethical complexity, or significant organizational consequences (Jarrahi, 2018; Rai, 2020). Human judgment remains necessary for interpreting algorithmic outputs, resolving ambiguity, balancing competing objectives, and evaluating recommended actions. Nevertheless, an AI system need not possess formal authority to influence which alternatives become visible, how risks are evaluated, or which options receive priority. The resulting governance problem is therefore a growing separation between the distribution of decision influence and the formal allocation of accountability. Managers and organizations may remain answerable for decisions whose informational framing, prediction, prioritization, or recommended action has been materially shaped by algorithmic systems. This challenges both performance measurement systems centred on identifiable decision-makers and accountability frameworks connecting authority, responsibility, traceability, and answerability. The theoretical challenge is not to redefine AI as an accountable organizational actor, but to explain how accountability can remain effective when decisions emerge through hybrid processes characterized by distributed decision influence and concentrated formal accountability.

ONLY FOR REVIEW

Existing Theories Become Insufficient

The preceding sections reveal a theoretical tension at the intersection of performance measurement, accountability, and AI-augmented decision-making. Performance measurement explains how organizations evaluate managerial actions and align behaviour with strategic objectives (Simons, 1995; Ferreira & Otley, 2009; Franco-Santos et al., 2012); accountability theory connects authority, responsibility, answerability, and legitimacy (Roberts, 1991; Sinclair, 1995; Bovens, 2007); and AI research examines interactions between human judgment and algorithmic capabilities (Faraj, Pachidi, & Sayegh, 2018; Dellermann et al., 2019; Rai, 2020; Raisch & Krakowski, 2021). The theoretical difficulty emerges from the incompatibility of their underlying assumptions when considered together. Performance measurement assumes that managerial actions and outcomes remain sufficiently connected to permit evaluation and attribution of responsibility (Kaplan & Norton, 1996; Ferreira & Otley, 2009; Merchant & Van der Stede, 2017), while accountability theory presupposes identifiable actors capable of explaining and justifying their conduct (Roberts, 1991; Bovens, 2007). Together, these perspectives imply an alignment among decision authority, responsibility, performance evaluation, and answerability (Simons, 1995; Franco-Santos et al., 2012; Bovens, 2007; Merchant & Van der Stede, 2017). AI-augmented decision-making complicates this alignment. AI systems can materially influence decisions through forecasts, evaluations, prioritizations, and recommendations (Jarrahi, 2018; Rai, 2020), while their organizational use remains dependent on human judgment and oversight (Dellermann et al., 2019; Raisch & Krakowski, 2021). From a socio-technical perspective, decisions therefore emerge through interactions among human actors, algorithmic systems, routines, and data infrastructures (Faraj et al., 2018). Consequently, decision influence may become distributed without a corresponding distribution of formal decision authority or accountability. AI governance research addresses related concerns through transparency, explainability, human oversight, and accountability (Floridi & Cowls, 2019; Wieringa, 2020), while Grote, Parker, and Crowston (2026) emphasize alignment between control and accountability among stakeholders involved in AI design and organizational use. Yet the organizational-level problem remains: how can accountability operate when decision influence is distributed across human and algorithmic contributors while formal accountability remains human and organizational? The limitation is therefore integrative. What remains insufficiently explained is how performance measurement, accountability, and AI-enabled decision-making can remain aligned once decision processes become hybrid.

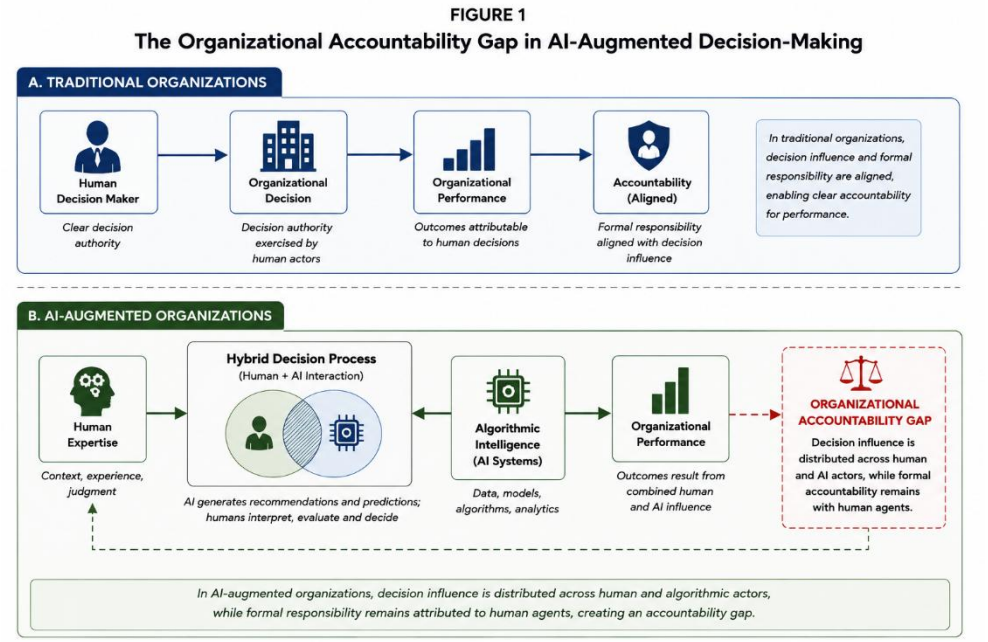
The Emerging Theoretical Puzzle

The preceding analysis identifies a fundamental inconsistency in the governance architecture of AI-augmented organizations. Performance measurement and accountability frameworks connect organizational outcomes to identifiable relationships among managerial action, authority, responsibility, and answerability, while AI introduces consequential algorithmic influence into decision processes. Decision influence can therefore become distributed across human and algorithmic contributors while formal decision authority and accountability remain concentrated in human and organizational actors. We conceptualize this condition as the Organizational Accountability Gap: a structural misalignment between the distribution of decision influence across human and algorithmic contributors and the

ONLY FOR REVIEW

formal allocation of accountability for organizational outcomes. The OAG does not imply that AI systems possess moral responsibility or should become accountable organizational actors. Rather, algorithmic systems may materially shape predictions, alternatives, priorities, and recommendations without acquiring corresponding answerability, while managers and organizations remain formally responsible for outcomes. The OAG thus represents a boundary condition arising when governance architectures developed around human-centred decision-making operate within increasingly hybrid decision processes. The theoretical question is not whether accountability should be transferred to AI, but how meaningful human and organizational accountability can be preserved when decision influence becomes distributed. This requires governance arrangements capable of making algorithmic influence transparent, traceable, contestable, and subject to meaningful organizational oversight. Existing theories therefore require extension rather than replacement. What remains theoretically underdeveloped is an integrative explanation of how decision influence, authority, responsibility, and performance evaluation can be coordinated under hybrid decision-making. The following section addresses this gap through the Theory of Hybrid Accountability, conceptualizing accountability as an organizational governance capability enacted through four interconnected mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. These mechanisms subsequently inform the Performance Accountability Architecture (PAA). As summarized in Figure 1, the misalignment between distributed decision influence and formal accountability constitutes the OAG and provides the theoretical foundation for the framework developed below.

Figure 1. *The Emergence of the Organizational Accountability Gap in AI-Augmented Organizations*



The figure illustrates the theoretical transition from traditional organizational governance to AI-augmented decision-making. In conventional organizations, decision authority, decision influence, organizational performance, and accountability are structurally aligned around human decision-makers. As Artificial Intelligence becomes increasingly embedded in organizational decision processes, decision influence is progressively distributed across human and algorithmic contributors, while formal accountability remains assigned to

human and organizational actors. This structural misalignment generates the Organizational Accountability Gap that motivates the development of the Theory of Hybrid Accountability.

Theory Development: Towards A Theory Of Hybrid Accountability

Reframing Accountability in AI-Augmented Organizations

The Organizational Accountability Gap requires reconsidering accountability in AI-augmented organizations. The relevant question is not whether AI systems should become accountable organizational actors, but how meaningful human and organizational accountability can be preserved when consequential decision influence is distributed across human and algorithmic contributors. This requires extending accountability from an exclusively actor-centred conception toward an organizational governance capability embedded throughout hybrid decision processes. The distinction between decision influence and responsibility provides the conceptual foundation for this reframing. Responsibility refers to the institutionally assigned obligation to explain, justify, and assume the consequences of organizational decisions, whereas decision influence refers to the capacity of human or algorithmic contributions to materially shape the information, alternatives, evaluations, or recommendations through which decisions are produced. In AI-augmented organizations, these dimensions may diverge: managers retain formal responsibility even when algorithmic outputs materially influence their decisions. This divergence changes the organizational requirements of accountability without transferring responsibility to AI. Organizations instead require governance arrangements capable of making hybrid decision processes visible, traceable, and governable. Accountability consequently concerns not only who is answerable for an outcome, but also how the decision process producing that outcome is structured and governed. Accountability should therefore be understood not merely as an ex post relationship, but as a governance capability embedded throughout the decision process. This capability depends upon four interconnected mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. Together, these mechanisms preserve human and organizational accountability while recognizing consequential algorithmic influence.

This process-oriented reframing provides the conceptual foundation for the Theory of Hybrid Accountability. Rather than replacing established accountability principles, Hybrid Accountability extends their application to organizational environments characterized by distributed decision influence.

Defining Hybrid Accountability

The Organizational Accountability Gap requires a conceptualization of accountability capable of explaining governance when decision influence is distributed across human and algorithmic contributors while formal responsibility remains human and organizational. Hybrid decision processes therefore require arrangements that preserve answerability while governing consequential algorithmic contributions.

We define Hybrid Accountability as an organizational governance capability through which algorithmic transparency, human oversight, decision authority, and responsibility

ONLINE REVIEW

allocation are coordinated to preserve meaningful accountability across hybrid decision processes characterized by distributed decision influence.

This definition establishes three properties. First, Hybrid Accountability is an organizational governance capability, rather than an attribute of an individual manager or technological system. Second, its object is the hybrid decision process, rather than the human or AI component considered independently. Third, its purpose is to preserve meaningful human and organizational accountability by governing the relationship between distributed decision influence and formally allocated authority and responsibility.

Hybrid Accountability extends rather than rejects established accountability theory. Conventional approaches conceptualize accountability as a relationship in which actors explain and justify their conduct before a forum capable of questioning, evaluating, and potentially sanctioning it (Bovens, 2007). Hybrid Accountability preserves answerability and responsibility while recognizing that organizational decisions increasingly incorporate consequential algorithmic contributions. It therefore concerns both who remains answerable and how the decision was produced, governed, and supervised.

Importantly, Hybrid Accountability does not transfer legal, ethical, or moral responsibility to AI. Human and organizational actors retain formal responsibility for decisions and their consequences. Its theoretical novelty lies instead in recognizing that meaningful human accountability increasingly depends on making algorithmic influence visible, traceable, governable, and subject to appropriate oversight. Thus, decision influence may be distributed across human and algorithmic contributors while responsibility and answerability remain institutionally anchored in human and organizational actors.

Hybrid Accountability also shifts attention from isolated decisions toward the governance of the decision process. Accountability operates across data generation, algorithmic processing and recommendations, managerial interpretation, implementation, and performance evaluation. It therefore encompasses ex ante, ongoing, and ex post governance, rather than exclusively retrospective evaluation.

Four characteristics further define the construct. Hybrid Accountability is process-embedded, because it operates throughout hybrid decision processes; relational, because it depends on interactions among organizational actors, algorithmic contributions, routines, and governance mechanisms; organizational, because it extends beyond individual managers while preserving identifiable human responsibility; and dynamic, because governance arrangements must adapt as technologies, organizational structures, and institutional requirements evolve.

These characteristics distinguish Hybrid Accountability from algorithmic accountability, Responsible AI, explainable AI, and human-in-the-loop governance. While these approaches primarily address the design, transparency, fairness, explainability, or supervision of AI systems, Hybrid Accountability focuses on the accountability of organizations operating through hybrid human–AI decision processes. The unit of analysis therefore shifts from the algorithm to the organizational governance system through which human judgment and algorithmic contributions interact.

This distinction also extends performance measurement. Traditional Performance Measurement Systems connect organizational objectives, managerial actions, indicators, and outcomes. Under distributed decision influence, performance evaluation must also consider the integrity, transparency, traceability, and governance quality of the hybrid processes producing those outcomes. Hybrid Accountability therefore provides a theoretical bridge

UNDER PEER REVIEW

among accountability theory, organizational governance, performance measurement, and AI-augmented decision-making.

The construct is enacted through four interconnected mechanisms—algorithmic transparency, human oversight, decision authority, and responsibility allocation—which provide the foundation for the subsequent Performance Accountability Architecture (PAA). Table 1 summarizes the conceptual distinction between Hybrid Accountability and traditional accountability models.

Table 1.

Dimension	Traditional Accountability	Hybrid Accountability
Decision architecture	Human-centred decision process	Hybrid human–AI decision process
Source of decision influence	Human judgment	Human judgment + algorithmic contributions
Formal decision authority	Human actors	Designated human actors within organizational authority structures
Accountability focus	Individual actor	Organizational governance capability
Unit of analysis	Individual decision	Hybrid decision process
Performance evaluation	Organizational outcomes	Outcomes + decision-process governance
Governance logic	Predominantly ex post accountability	Ex ante, ongoing, and ex post governance
Role of AI	Decision-support tool	Source of consequential decision influence
Transparency	Managerial reporting and explanation	Algorithmic traceability + managerial explanation
Oversight	Managerial supervision	Meaningful human oversight of algorithmic contributions
Responsibility	Individual managerial responsibility	Explicit organizational responsibility allocation
Performance Measurement System	Performance measurement and control system	Performance and accountability governance infrastructure
Organizational objective	Performance control	Performance control + meaningful organizational accountability

The Generative Mechanism of Hybrid Accountability

Having conceptualized Hybrid Accountability as an organizational governance capability, the next theoretical challenge is to explain how and under what conditions it emerges. AI adoption is an enabling rather than sufficient condition: Hybrid Accountability becomes relevant when algorithmic systems acquire consequential influence over organizational decisions while formal authority and accountability remain human and organizational. The generative process begins as algorithmic predictions, classifications, rankings, risk assessments, or recommendations materially shape managerial judgment. Decision influence consequently becomes distributed across human and algorithmic

ONLY FOR REVIEW

contributors without a corresponding redistribution of formal decision authority. As this divergence increases, the distribution of decision influence becomes progressively misaligned with the allocation of formal accountability, producing the Organizational Accountability Gap. The OAG does not automatically produce Hybrid Accountability. Rather, it generates governance pressure to adapt existing accountability arrangements. Because responsibility cannot simply be transferred to AI systems, organizations must develop mechanisms that make algorithmic influence visible and traceable, preserve meaningful human oversight, clarify formal decision authority, and allocate responsibility across relevant organizational roles. Hybrid Accountability emerges through this coordinated governance adaptation, reconnecting distributed decision influence with meaningful human and organizational accountability.

This adaptation also shifts accountability from predominantly actor-centred and ex post arrangements toward process-embedded governance. Accountability becomes embedded in how algorithmic outputs are generated, interpreted, questioned, authorized, implemented, and evaluated, thereby preserving human responsibility throughout hybrid decision processes. The process is neither automatic nor uniform. Governance pressure is likely to increase with the degree of algorithmic influence, the consequential importance of decisions, and the difficulty of reconstructing algorithmic contributions. Where AI remains limited to low-consequence analytical support and managers retain substantial independent control, conventional accountability arrangements may remain comparatively adequate. Governance maturity, regulatory requirements, managerial capabilities, organizational risk exposure, and available oversight mechanisms may further shape organizations' capacity to respond to the OAG. Accordingly, the generative mechanism involves three interrelated processes: (1) distributed decision influence; (2) persistent human and organizational authority and accountability; and (3) governance adaptation aimed at reconnecting these dimensions. The resulting sequence can therefore be expressed as: distributed decision influence → Organizational Accountability Gap → governance pressure and adaptation → Hybrid Accountability. The following section specifies the four mechanisms through which this capability is enacted: algorithmic transparency, human oversight, decision authority, and responsibility allocation.

Governance Mechanisms of Hybrid Accountability

The generative mechanism explains why Hybrid Accountability emerges; the next question is how this capability is enacted. We argue that Hybrid Accountability operates through four mutually reinforcing governance mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. These mechanisms are not independent attributes or sequential stages, but an integrated governance architecture through which organizations reconnect distributed decision influence with formal human and organizational accountability.

Algorithmic Transparency

Algorithmic transparency concerns the visibility and traceability of algorithmic influence within organizational decision processes. As AI generates forecasts, classifications, risk assessments, and recommendations, meaningful accountability requires organizations to understand how these outputs enter and influence managerial decisions. Within Hybrid

ONLY FOR REVIEW

Accountability, transparency extends beyond technical explainability. It concerns the organizational visibility of the information, assumptions, limitations, uncertainty, and decision relevance associated with algorithmic contributions. Its purpose is not to make every computational operation intelligible, but to provide sufficient information for decision-makers to evaluate algorithmic outputs within their organizational context.

Transparency therefore makes decision influence traceable, enabling managers to question algorithmic recommendations and explain how they contributed to organizational decisions. It does not redistribute responsibility but establishes the informational conditions for meaningful human oversight.

Human Oversight

Transparency alone cannot preserve accountability. Human oversight refers to the capacity of designated human actors to evaluate, question, contest, modify, or override algorithmic contributions when they are unreliable, inappropriate, insufficiently contextualized, or inconsistent with organizational objectives.The intensity of oversight may vary with the strategic significance, uncertainty, reversibility, and consequences of a decision. What matters is that algorithmic influence remains subject to meaningful human judgment rather than merely formal human approval. Effective oversight therefore requires both authority and practical capability to interrogate algorithmic outputs, incorporate contextual knowledge, resolve ambiguity, and intervene when necessary. Human oversight consequently converts algorithmic transparency into organizational judgment.

Decision Authority

Decision authority concerns the explicit location of formal authority within hybrid decision processes. AI weakens the traditional correspondence between decision influence and authority because algorithmic systems may materially influence choices while managers retain formal authority. Hybrid Accountability therefore distinguishes the two dimensions: decision influence may be distributed across human and algorithmic contributors, while formal decision authority remains vested in designated human and organizational actors. Organizations must consequently specify not only who possesses authority, but when human actors are expected to review, authorize, modify, escalate, or reject AI-supported recommendations. Decision authority thereby preserves the institutional connection between organizational judgment and human answerability without attributing authority to AI systems.

Responsibility Allocation

Responsibility allocation concerns how accountability-related obligations associated with AI-enabled decision processes are distributed among identifiable human and organizational actors. Responsibility is therefore differentiated rather than transferred. Managers remain answerable for organizational decisions, while other actors may hold responsibilities for model and data quality, monitoring, compliance, auditing, or the integrity of the supporting decision infrastructure.Because hybrid decision processes extend beyond final authorization, relevant responsibilities may span data preparation, system design, implementation, validation, monitoring, interpretation, and organizational use. Hybrid Accountability coordinates these differentiated responsibilities without diluting final human

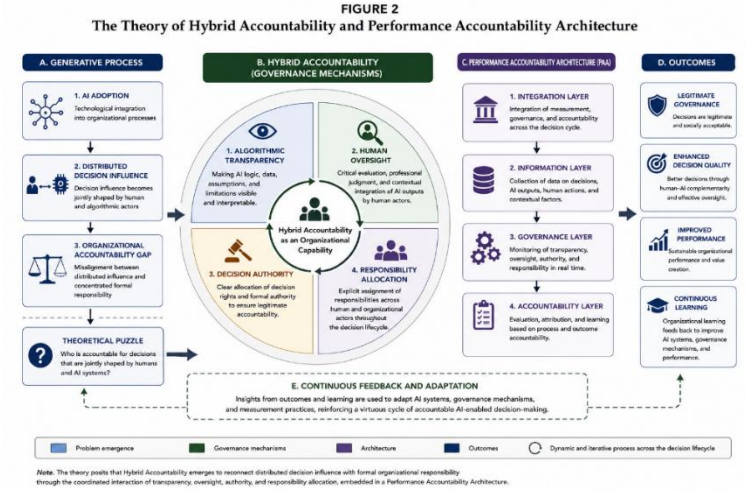
ONLY FOR REVIEW

and organizational answerability, transforming accountability from an isolated managerial obligation into a coordinated organizational arrangement.

Integrating the Governance Mechanisms

Together, the four mechanisms constitute the internal governance architecture of Hybrid Accountability. Transparency makes algorithmic influence visible; oversight enables its critical evaluation; decision authority establishes where legitimate organizational judgment resides; and responsibility allocation distributes accountability-related obligations across relevant human and organizational actors. Their theoretical significance lies in their interdependence. Transparency without oversight does not ensure intervention; oversight without decision authority leaves final judgment ambiguous; authority without responsibility allocation leaves the broader governance infrastructure unclear; and responsibility allocation without transparency and oversight may assign duties without enabling their effective exercise. Hybrid Accountability therefore emerges from the coordinated interaction of these mechanisms rather than their additive presence. Transparency enables oversight; oversight supports meaningful decision authority; authority anchors answerability; and responsibility allocation sustains the organizational infrastructure through which these mechanisms operate. Their interaction reconnects distributed decision influence with formal human and organizational accountability without attributing authority or answerability to AI. This architecture also connects Hybrid Accountability to performance measurement. If outcomes emerge through hybrid decision processes, Performance Measurement Systems must consider not only outcomes and managerial behaviour but also whether algorithmic influence was transparent, oversight meaningful, decision authority appropriately exercised, and responsibilities clearly allocated. The four mechanisms therefore provide the organizational foundation for the Performance Accountability Architecture (PAA) developed in the following section.

Figure 2. *The Theory of Hybrid Accountability and the Performance Accountability Architecture*



The figure presents the proposed theoretical framework developed in this study. AI adoption progressively increases distributed decision influence, generating an organizational accountability gap through the decoupling of decision influence and formal responsibility. Hybrid Accountability emerges as an organizational governance capability enacted through four mutually reinforcing governance mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. These mechanisms are institutionalized

FOR REVIEW

through the Performance Accountability Architecture, which enables adaptive performance measurement and supports effective organizational governance in AI-augmented organizations.

Performance Accountability Architecture

The Theory of Hybrid Accountability requires reconsidering how accountability is connected to performance measurement and control. Traditional Performance Measurement Systems (PMS) monitor managerial behaviour, evaluate organizational outcomes, and align actions with strategic objectives (Kaplan & Norton, 1996; Ferreira & Otley, 2009; Franco-Santos et al., 2012). However, when decision influence becomes distributed across human and algorithmic contributors, evaluating outcomes and managerial behaviour alone provides an incomplete representation of how organizational performance is generated. Organizational control must therefore consider not only what outcomes are achieved, but how the hybrid decisions producing them are governed. To address this requirement, we introduce the Performance Accountability Architecture (PAA), defined as the organizational governance architecture through which the mechanisms of Hybrid Accountability are embedded within performance measurement and control, enabling organizations to evaluate both organizational outcomes and the governance integrity of the hybrid decision processes through which those outcomes are produced. PAA does not replace conventional performance measurement. Rather, it extends organizational control by incorporating governance conditions associated with algorithmic transparency, human oversight, decision authority, and responsibility allocation. Its central distinction is therefore between performance outcomes and decision-process governance. PAA comprises three interdependent governance layers: decision generation, decision governance, and performance accountability. These are analytically distinct but continuously interacting components of the organizational decision cycle. Decision generation concerns how data and algorithmic processing produce decision-relevant predictions, classifications, and recommendations. Governance at this layer focuses on the visibility and traceability of algorithmic influence, including data quality, model limitations, uncertainty, explainability, and the ability to reconstruct how algorithmic outputs entered the decision process. Decision governance concerns how algorithmic contributions are evaluated, contested, contextualized, and translated into authorized organizational decisions. Human oversight ensures that algorithmic outputs remain subject to critical evaluation, while decision authority identifies the human and organizational actors responsible for authorizing consequential decisions. This layer therefore transforms algorithmic influence into organizational action without conflating influence with formal authority. Performance accountability evaluates organizational outcomes together with the accountability conditions surrounding the hybrid decision process. It considers whether algorithmic influence remained traceable, appropriate oversight occurred, decision authority was meaningfully exercised, and responsibilities were clearly allocated. Performance evaluation consequently incorporates both outcomes and the governance quality of the processes through which they were produced. The three layers form a recursive governance architecture. Performance accountability feeds information back into subsequent decision generation and governance by revealing weaknesses in transparency, oversight, authority, or responsibility allocation. Changes in AI capabilities and organizational processes may likewise require adjustments in performance and accountability criteria. Conventional financial, operational, strategic, and behavioural measures therefore remain relevant but are complemented by governance-oriented information concerning

ONLY FOR REVIEW

algorithmic traceability, oversight effectiveness, decision authority, and responsibility clarity. Hybrid Accountability and PAA are thus complementary but distinct: Hybrid Accountability constitutes the organizational governance capability, whereas PAA constitutes the architecture through which that capability is institutionalized within performance measurement and control. The theoretical progression can therefore be summarized as: Distributed decision influence → Organizational Accountability Gap → governance pressure and adaptation → Hybrid Accountability → governance mechanisms → Performance Accountability Architecture. PAA consequently extends performance measurement from a predominantly outcome-oriented control function toward an integrated governance infrastructure connecting organizational performance with the transparency, oversight, authority, and responsibility conditions under which hybrid decisions are produced. This architecture provides the basis for the theoretical propositions developed below.

Theoretical Propositions

The Emergence of Hybrid Accountability

AI adoption alone does not transform accountability. The relevant condition is the extent to which AI acquires consequential influence over organizational decisions while formal authority and accountability remain human and organizational. As this asymmetry increases, the Organizational Accountability Gap intensifies, creating pressure for governance adaptation. Where AI remains limited to analytical support, conventional accountability may remain adequate; where algorithmic outputs materially shape organizational decisions, Hybrid Accountability becomes increasingly necessary.

Proposition 1. The greater the consequential influence of AI on organizational decision processes, the greater the Organizational Accountability Gap and, consequently, the stronger the organizational need for Hybrid Accountability.

Governance Mechanisms of Hybrid Accountability

Algorithmic transparency enables meaningful oversight by making the assumptions, limitations, uncertainty, and decision relevance of AI-generated outputs more visible and traceable. Its governance value therefore depends on organizational actors being able to use this information to evaluate and contest algorithmic contributions.

Proposition 2. Higher levels of algorithmic transparency increase the effectiveness of human oversight by improving the traceability, interpretability, and contestability of algorithmic contributions to organizational decisions.

Human oversight connects algorithmic influence with organizational judgment. Where managers can critically evaluate AI-generated recommendations, formal human involvement becomes substantively meaningful rather than merely procedural.

Proposition 3. More effective human oversight strengthens Hybrid Accountability by increasing the extent to which formal human decision authority is exercised through meaningful organizational judgment rather than nominal approval of algorithmic recommendations.

ONLY FOR REVIEW

Decision authority further preserves accountability by specifying where legitimate organizational authority resides and when human authorization is required. Distributed decision influence therefore does not imply distributed formal authority.

Proposition 4. Greater clarity in the allocation and exercise of formal decision authority strengthens the relationship between human oversight and organizational accountability within AI-augmented decision processes.

Responsibility allocation extends accountability across the multiple human and organizational roles involved in data generation, system design, validation, monitoring, interpretation, authorization, auditing, and organizational use. Clear allocation reduces ambiguity without diluting final human and organizational answerability.

Proposition 5. Greater clarity in responsibility allocation across the AI-enabled decision process strengthens Hybrid Accountability by reducing ambiguity concerning the obligations of organizational actors while preserving identifiable human and organizational answerability.

From Hybrid Accountability to Performance Accountability Architecture

As Hybrid Accountability develops, performance evaluation must increasingly consider not only organizational outcomes but also the governance conditions under which hybrid decisions are produced. This includes algorithmic traceability, meaningful human oversight, the exercise of decision authority, and responsibility allocation, thereby connecting Hybrid Accountability to the Performance Accountability Architecture.

Proposition 6. Higher levels of Hybrid Accountability are associated with greater integration of governance-oriented criteria into Performance Measurement Systems, thereby promoting the development of Performance Accountability Architectures.

Taken together, the propositions specify the explanatory structure of the theory. Proposition 1 explains why Hybrid Accountability becomes necessary; Propositions 2–5 specify how its governance mechanisms operate; and Proposition 6 connects Hybrid Accountability to performance measurement and PAA. The resulting framework therefore moves from distributed decision influence and the Organizational Accountability Gap to governance adaptation, Hybrid Accountability, and its institutionalization within performance measurement and control. These propositions provide a foundation for empirical examination. Future research may operationalize algorithmic decision influence, transparency, oversight effectiveness, authority clarity, responsibility allocation, and governance-oriented performance measurement across organizational and technological contexts. Qualitative and longitudinal research may further examine how these relationships evolve as AI becomes more deeply embedded in organizational decision processes.

UNDER PEER REVIEW

Discussion And Implications

Theoretical Contributions

The Theory of Hybrid Accountability integrates accountability theory, performance measurement, and AI governance to explain how meaningful accountability can be preserved when decision influence becomes distributed across human and algorithmic contributors while formal authority and accountability remain human and organizational. The study makes three principal theoretical contributions. First, it extends organizational accountability theory by distinguishing decision influence from formal accountability and conceptualizing accountability as an organizational governance capability embedded within hybrid decision processes. Traditional accountability theories emphasize answerability, responsibility, explanation, and justification among identifiable organizational actors (Roberts, 1991; Sinclair, 1995; Bovens, 2007). Hybrid decision-making introduces a boundary condition because consequential decision influence may originate partly from algorithmic systems possessing neither formal authority nor institutional answerability. The Organizational Accountability Gap captures this structural misalignment. Hybrid Accountability addresses it without transferring accountability to AI: answerability remains human and organizational, while its preservation increasingly depends on process-embedded governance through algorithmic transparency, human oversight, decision authority, and responsibility allocation. Accountability thus remains human-centred in its institutional location while becoming hybrid in the governance architecture required to sustain it. Second, the study extends performance measurement theory by incorporating the governance quality of hybrid decision processes into organizational control. Traditional Performance Measurement Systems connect strategic objectives, managerial actions, indicators, and outcomes through predominantly human-centred control architectures (Kaplan & Norton, 1996; Ferreira & Otley, 2009; Franco-Santos et al., 2012). The Performance Accountability Architecture (PAA) extends this logic by integrating performance outcomes with decision-process governance. Conventional financial, operational, strategic, and behavioural measures remain relevant but are complemented by governance-oriented dimensions concerning algorithmic traceability, oversight effectiveness, decision authority, and responsibility clarity. The object of performance evaluation therefore expands from outcomes and managerial behaviour to the governance conditions producing those outcomes, while control extends across ex ante, ongoing, and ex post accountability. Third, the study broadens AI governance research by shifting the unit of analysis from AI systems considered in isolation toward organizations operating through AI-augmented decision processes. Algorithmic transparency, explainability, fairness, auditing, and human oversight remain important, but organizational accountability requires these mechanisms to be connected to structures of authority and responsibility. Hybrid Accountability therefore explains how transparency enables oversight, how oversight supports meaningful human decision authority, and how responsibility is allocated across the organizational roles involved in AI-enabled decisions. AI governance consequently becomes part of a broader organizational governance architecture rather than exclusively a technology-centred concern. Collectively, these contributions establish a common theoretical logic: distributed decision influence generates the Organizational Accountability Gap; the gap creates governance pressure and adaptation; Hybrid Accountability emerges as an organizational governance capability enacted through four

UNDER PEER REVIEW

interconnected mechanisms; and these mechanisms are institutionalized within performance measurement and control through the PAA. The broader contribution is therefore to explain how human-centred accountability can be preserved within non-exclusively-human decision processes. Human and organizational actors remain answerable, but meaningful accountability increasingly depends on organizations' capacity to make algorithmic influence traceable, subject it to oversight, preserve legitimate decision authority, and clearly allocate responsibilities throughout hybrid decision processes.

Contributions to Performance Measurement Research

Beyond accountability theory, the Theory of Hybrid Accountability extends the conceptual domain of Performance Measurement Systems (PMS). Performance measurement has evolved from financially oriented control toward multidimensional frameworks supporting strategic alignment, organizational learning, behavioural coordination, and value creation (Neely, Gregory, & Platts, 1995; Kaplan & Norton, 1996; Ferreira & Otley, 2009; Franco-Santos et al., 2012). Yet these frameworks largely assume that organizational performance can ultimately be connected to identifiable human decision-makers. AI-augmented decision-making challenges this assumption because outcomes may increasingly emerge from interactions between human judgment and algorithmic contributions. The framework therefore extends performance measurement along three dimensions: what is measured, when governance-oriented measurement operates, and how performance measurement relates to accountability. First, it broadens the object of performance measurement. Conventional financial, operational, strategic, and behavioural indicators remain essential, but under distributed decision influence they are complemented by the governance integrity of hybrid decision processes. The Performance Accountability Architecture (PAA) therefore incorporates algorithmic traceability, oversight effectiveness, the exercise of decision authority, and responsibility clarity. Performance measurement consequently evaluates both what was achieved and under what governance conditions the decisions producing those outcomes were generated and authorized. Second, the framework extends the temporal logic of performance measurement and control. Management control traditionally combines ex post evaluation with planning, feedback, learning, and behavioural control (Simons, 1995; Merchant & Van der Stede, 2017). Hybrid Accountability extends these functions across ex ante, ongoing, and ex post governance. Ex ante governance establishes transparency, authority, and responsibility conditions; ongoing governance supports monitoring and contestation of algorithmic contributions; and ex post evaluation connects outcomes with managerial actions and the governance quality of the decision process. PMS thereby acquire a more explicit process-governance function alongside their established evaluative and control functions. Third, the framework redefines the relationship between performance measurement and accountability. PMS traditionally provide information through which managerial performance can be evaluated and actors held answerable. Under Hybrid Accountability, accountability requirements also determine which dimensions of hybrid decision processes become objects of measurement and control. This creates a reciprocal relationship: PMS provide information supporting accountability, while accountability requirements shape what PMS must represent. PAA formalizes this relationship within a common governance architecture. These extensions do not require replacing existing PMS or abandoning conventional financial, operational, strategic,

ONLY FOR REVIEW

behavioural, or Balanced Scorecard measures. Rather, their informational architecture is extended to represent whether performance was generated under conditions consistent with meaningful human accountability. AI therefore changes not only the information available for performance measurement but also the relationship among decision influence, performance attribution, and organizational accountability. Overall, the Theory of Hybrid Accountability extends performance measurement in terms of what is measured, when governance-oriented measurement operates, and how measurement supports accountability. PAA integrates these dimensions by treating organizational outcomes and the governance quality of hybrid decision processes as complementary objects of organizational control. Table 2 summarizes how the four mechanisms of Hybrid Accountability translate into specific extensions of performance measurement and their expected organizational implications.

Table 2.

Governance Mechanism	Governance Objective	Performance Measurement Implication	Expected Governance Implication
Algorithmic Transparency	Make consequential AI influence visible and traceable	Indicators of algorithmic traceability and decision-process transparency	Enhanced traceability and governance transparency
Human Oversight	Preserve meaningful human evaluation and intervention	Indicators of oversight effectiveness and human review	More meaningful human oversight
Decision Authority	Clarify the location and exercise of formal decision authority	Indicators of authorization, escalation, and decision rights	Clearer authorization and answerability
Responsibility Allocation	Establish explicit accountability-related responsibilities across organizational roles	Indicators of responsibility allocation and role clarity	Clearer responsibility attribution

Contributions to Organizational Governance and AI Research

The Theory of Hybrid Accountability contributes to organizational governance and AI research by shifting attention from the governance of AI technologies toward the governance of organizations operating through AI-augmented decision processes. AI governance research has primarily examined how algorithmic systems should be designed, monitored, and regulated to promote fairness, transparency, explainability, robustness, and regulatory compliance (Floridi & Cowls, 2019; Wieringa, 2020). Hybrid Accountability complements this perspective by asking how organizations should be governed when AI materially influences consequential decisions. This shift changes the unit of analysis from the AI system to the organizational governance architecture through which algorithmic contributions are interpreted, contested, authorized, implemented, and connected to human accountability. The framework extends AI governance in three respects. First, it conceptualizes governance mechanisms as an interdependent system. Algorithmic transparency alone cannot preserve accountability unless human actors can meaningfully evaluate algorithmic contributions,

ONLY FOR REVIEW

possess legitimate authority to intervene, and operate within clearly allocated responsibilities. The contribution therefore lies not in introducing transparency, oversight, authority, or responsibility individually, but in explaining why their coordinated interaction becomes necessary when decision influence is distributed while formal accountability remains human and organizational. Second, the theory distinguishes technological trustworthiness from organizational accountability. Explainability, transparency, robustness, and other properties of trustworthy AI remain important, but a trustworthy AI system does not necessarily produce an accountable organizational decision process. Accountability also depends on how algorithmic contributions are evaluated, authorized, acted upon, and connected to identifiable human and organizational responsibilities. Trustworthy AI and accountable organizations are therefore related but analytically distinct governance objectives. Third, the theory conceptualizes AI-enabled organizational change as a governance transformation as well as a technological and strategic transformation. As algorithmic influence increases, organizations must adapt not only technologies and decision routines but also structures of oversight, authority, and responsibility. The Organizational Accountability Gap captures the governance pressure arising when these structures fail to evolve alongside changes in decision influence. Hybrid Accountability therefore identifies accountability architecture as a dimension of organizational adaptation to AI. This perspective remains consistent with a socio-technical understanding of AI-enabled decision-making while maintaining a clear distinction between algorithmic decision influence and formal organizational authority. AI systems may materially influence decisions without acquiring organizational authority or accountability. Formal answerability remains human and organizational, while its effectiveness increasingly depends on mechanisms that make algorithmic influence traceable, subject it to meaningful oversight, preserve identifiable decision authority, and allocate responsibilities throughout the hybrid decision process. Overall, Hybrid Accountability broadens AI governance from a predominantly technology-centred perspective toward an organization-centred governance perspective without replacing existing principles of responsible and trustworthy AI. It embeds technological governance within an organizational framework explaining how algorithmic influence interacts with human authority and institutional accountability.

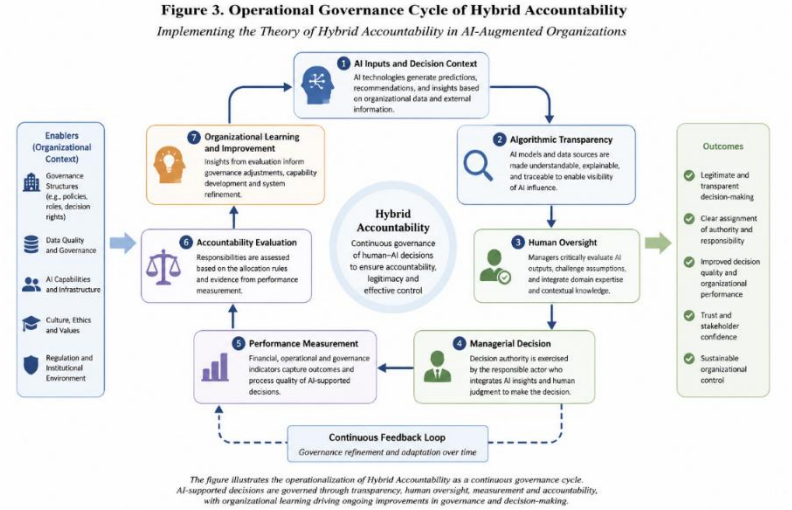
Managerial Implications

The Theory of Hybrid Accountability provides practical implications for organizations in which AI increasingly influences consequential decisions. In addition to technological capability, data quality, regulatory compliance, and responsible AI principles, organizations must govern the decision processes through which algorithmic contributions acquire organizational influence. First, organizations should explicitly define decision authority by identifying where AI materially influences organizational choices and specifying which human actors are authorized to review, modify, escalate, reject, or approve AI-supported recommendations. The objective is to ensure that increasing algorithmic influence does not reduce formal human authority to nominal approval. Second, organizations should institutionalize meaningful human oversight. Human involvement alone does not ensure accountability; designated actors require sufficient information, competence, time, and authority to critically evaluate algorithmic contributions. Oversight intensity should reflect the consequentiality, uncertainty, reversibility, and organizational significance of the

ONLY FOR REVIEW

decision. Third, algorithmic transparency and decision traceability should enable managers to reconstruct how consequential algorithmic contributions entered the decision process. Relevant inputs, assumptions, recommendations, limitations, and human interventions should therefore be sufficiently visible to support evaluation, justification, and subsequent review. Fourth, organizations should explicitly allocate responsibilities across the hybrid decision process. Responsibilities for data quality, system deployment, monitoring, interpretation, human review, authorization, implementation, auditing, and performance evaluation may be differentiated across organizational roles without transferring accountability to AI or diluting final human and organizational answerability. These principles also have implications for Performance Measurement Systems. Where algorithmic influence is consequential, conventional financial, operational, strategic, and behavioural indicators can be complemented by governance-oriented measures assessing whether algorithmic contributions were traceable, required oversight occurred, decision authority was meaningfully exercised, and responsibilities were clearly allocated. Such measures should follow a proportionate governance logic rather than a universal checklist: low-consequence applications may require limited controls, whereas strategic, high-impact, or difficult-to-reverse decisions may require stronger transparency, oversight, authorization, and documentation. Implementing Hybrid Accountability also requires cross-functional coordination among managerial, technological, accounting, legal, risk, audit, and compliance functions. PAA provides a framework for differentiating their responsibilities while integrating them within a coherent accountability architecture. Finally, Hybrid Accountability should operate as a recursive governance cycle rather than a one-time compliance exercise. Performance evaluation can reveal weaknesses in transparency, oversight, authority, or responsibility allocation and feed these findings back into subsequent governance design. Organizations should therefore reassess accountability arrangements as AI systems, organizational uses, and patterns of decision influence evolve.

Figure 3. *Operational Governance Cycle of Hybrid Accountability*



The figure illustrates how organizations operationalize the Theory of Hybrid Accountability through a continuous governance cycle. AI-supported decision-making is governed through algorithmic transparency, human oversight, managerial decision authority, performance measurement, and accountability evaluation. Organizational learning provides continuous feedback, enabling governance mechanisms and Performance Measurement Systems to evolve as AI capabilities and organizational contexts change.

FOR REVIEW

Limitations and Future Research

As a conceptual and theory-building study, this paper establishes a theoretical foundation for accountability in AI-augmented organizations while identifying boundary conditions requiring further theoretical and empirical examination. The Theory of Hybrid Accountability and the Performance Accountability Architecture (PAA) are intended to explain how meaningful human and organizational accountability can be preserved under distributed decision influence, but their explanatory scope requires validation across different organizational, institutional, and technological contexts. A first limitation concerns organizational and institutional boundary conditions. The framework is developed at a general organizational level and does not specify how Hybrid Accountability varies across industries, regulatory environments, ownership structures, organizational forms, or degrees of decision centralization. Future research should examine whether different institutional settings require different configurations of the four governance mechanisms. Second, the framework does not fully capture heterogeneity in AI technologies and their organizational roles. Predictive systems, generative AI, recommendation systems, and more autonomous applications may produce different forms and degrees of consequential decision influence. Future research should therefore investigate how these differences affect requirements for transparency, oversight, decision authority, and responsibility allocation. Third, Hybrid Accountability is inherently dynamic. Governance arrangements may emerge before, during, or after AI integration and evolve as technologies, organizational routines, regulatory requirements, and reliance on algorithmic recommendations change. Longitudinal research could examine how the Organizational Accountability Gap and corresponding governance arrangements develop across different stages of AI integration. Fourth, the Organizational Accountability Gap, Hybrid Accountability, and PAA are introduced as theoretical constructs rather than established empirical measures. Future research should develop and validate appropriate operationalizations while preserving two central distinctions: AI adoption should not be equated with consequential algorithmic decision influence, and formal human involvement should not be equated with meaningful human oversight. These limitations provide several opportunities for empirical research. Quantitative studies could test the proposed relationships among algorithmic influence, the Organizational Accountability Gap, the four governance mechanisms, Hybrid Accountability, and PAA. Qualitative, comparative, and longitudinal studies could investigate how accountability arrangements are constructed, negotiated, and adapted across organizational contexts. Research could also examine how governance-oriented indicators concerning algorithmic traceability, oversight effectiveness, decision authority, and responsibility clarity can be integrated into existing Performance Measurement Systems. Future research should further examine the organizational consequences and potential trade-offs of Hybrid Accountability. Stronger governance arrangements may support legitimacy, stakeholder trust, organizational learning, resilience, and decision quality, but may also generate coordination costs, slower decision processes, additional documentation, or excessive control when disproportionate to algorithmic influence. Identifying the conditions under which different governance configurations generate benefits or burdens therefore represents an important empirical question. Finally, further conceptual research could clarify the relationship between Hybrid Accountability and adjacent perspectives in management accounting, organizational governance, information systems, responsible AI, and socio-technical research. Overall, the

UNDER PEER REVIEW

framework should be understood as a theory-building platform rather than a closed explanatory model. Future research can refine its boundary conditions, develop empirical measures, test its propositions, and examine how different configurations of Hybrid Accountability operate across organizational and institutional contexts.

Conclusion

The increasing integration of Artificial Intelligence into organizational decision-making is altering the conditions under which organizations exercise control, evaluate performance, and assign accountability. As algorithmic systems acquire consequential influence, decision influence may become distributed across human and algorithmic contributors while formal decision authority and accountability remain human and organizational. This paper conceptualizes this tension as the Organizational Accountability Gap: a structural misalignment between the distribution of decision influence across human and algorithmic contributors and the formal allocation of accountability for organizational outcomes. Addressing this gap does not require transferring accountability to AI, but governing the relationship among algorithmic influence, human judgment, formal authority, and organizational responsibility. The paper therefore develops the Theory of Hybrid Accountability, conceptualizing accountability as an organizational governance capability enacted through four interconnected mechanisms: algorithmic transparency, human oversight, decision authority, and responsibility allocation. AI systems may materially influence organizational decisions, but formal authority and accountability remain human and organizational. Hybrid Accountability explains how meaningful human accountability can be preserved under these conditions. The study further introduces the Performance Accountability Architecture (PAA) as the architecture through which these mechanisms are embedded within performance measurement and control. PAA extends Performance Measurement Systems beyond outcomes and managerial behaviour to incorporate the governance integrity of hybrid decision processes, connecting accountability to ex ante, ongoing, and ex post governance. The theoretical progression can therefore be summarized as: distributed decision influence → Organizational Accountability Gap → governance pressure and adaptation → Hybrid Accountability → governance mechanisms → Performance Accountability Architecture. This framework extends accountability theory by distinguishing decision influence from formal accountability; performance measurement research by incorporating decision-process governance into organizational evaluation; and AI governance by shifting attention from the governance of technological systems toward the governance of organizations operating through AI-augmented decision processes. As a theory-building framework, Hybrid Accountability also establishes an agenda for empirical research. Future studies can refine its boundary conditions, operationalize its constructs, test the proposed relationships, and examine the benefits and costs of different governance configurations. Ultimately, the organizational challenge created by AI is not simply to develop or adopt more capable technologies, but to ensure that increasing algorithmic influence does not weaken the conditions under which decisions remain explainable, contestable, authorized, and attributable to responsible human actors. Human-centred accountability can be preserved within decision processes that are no longer exclusively

UNDER PEER REVIEW

human—but only when the organizational architecture governing those processes evolves alongside the technologies that influence them.

References

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal, 13*(4), 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering, 61*(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization, 28*(1), 62–70. https://doi.org/10.1016/j.infoandorg.2018.02.005

Ferreira, A., & Otley, D. (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research, 20*(4), 263–282. https://doi.org/10.1016/j.mar.2009.07.003

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.8cd550d1

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research, 23*(2), 79–119. https://doi.org/10.1016/j.mar.2012.04.001

Grote, G., Parker, S. K., & Crowston, K. (2026). Taming artificial intelligence: A theory of control–accountability alignment among AI developers and users. *Academy of Management Review, 51*(2), 278–299. https://doi.org/10.5465/amr.2023.0117

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action*. Harvard Business School Press.

Merchant, K. A., & Van der Stede, W. A. (2017). *Management control systems: Performance measurement, evaluation and incentives* (4th ed.). Pearson.

Messner, M. (2009). The limits of accountability. *Accounting, Organizations and Society, 34*(8), 918–938. https://doi.org/10.1016/j.aos.2009.07.003

Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management, 15*(4), 80–116. https://doi.org/10.1108/01443579510083622

Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society, 39*(4), 1871–1882. https://doi.org/10.1007/s00146-023-01635-y

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Roberts, J. (1991). The possibilities of accountability. *Accounting, Organizations and Society, 16*(4), 355–368. https://doi.org/10.1016/0361-3682(91)90027-C

Roberts, J. (2009). No one is perfect: The limits of transparency and an ethic for "intelligent" accountability. *Accounting, Organizations and Society, 34*(8), 957–970. https://doi.org/10.1016/j.aos.2009.04.005

Simons, R. (1995). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Harvard Business School Press.

ONLY FOR REVIEW

Sinclair, A. (1995). The chameleon of accountability: Forms and discourses. *Accounting, Organizations and Society, 20*(2–3), 219–237. https://doi.org/10.1016/0361-3682(93)E0003-Y

Wieringa, M. (2020). What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372833

ONLY FOR REVIEW