

Human-Artificial Intelligence Collaboration in Strategic Decision-Making: A Systematic Literature Review

Margret Yusuf Malgwi¹, Zakari Mohammed¹, Caleb Markus^{2*}

¹Department of Business Administration, Modibbo Adama University, Yola, Nigeria

²Department of Computer Science, Taraba State University, Jalingo, Nigeria

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100900123>

Received: 19 September 2026; Accepted: 24 September 2026; Published: 02 October 2026

ABSTRACT

Artificial intelligence is increasingly involved in strategic decision-making, yet its contribution depends on how decision authority is distributed between humans and machines. This systematic literature review critically examines AI delegation, where analytical or decision tasks are transferred to AI, and AI augmentation, where AI informs but does not replace human judgment. Guided by PRISMA 2020, the review screened 4,087 records and retained 25 studies for thematic synthesis. The evidence was interpreted through bounded rationality, algorithm aversion and appreciation, organization design, and the automation-augmentation paradox. Five themes emerged: the conditional value of delegation, the complementarity of augmentation, trust calibration and expertise, task structure and uncertainty, and organizational governance. The synthesis shows that delegation can improve speed, consistency, and analytical scale when problems are structured and feedback is reliable, but can reduce contestability and magnify model error when strategic ambiguity is high. Augmentation is more robust for uncertain, irreversible, politically sensitive, and novel strategic problems because human actors can contextualize predictions, challenge outputs, and integrate tacit knowledge. However, augmentation is not automatically superior because poorly calibrated trust, anchoring, automation bias, and weak AI literacy can make human oversight nominal rather than substantive. The review therefore rejects a universal choice between automation and augmentation and proposes a contingent collaboration framework in which decision quality depends on matching AI engagement mode to task characteristics, human expertise, and governance capacity. The study contributes a strategic-management-specific synthesis and identifies priorities for longitudinal, field-experimental, and generative-AI research.

Keywords: artificial intelligence; human-AI collaboration; strategic decision-making; AI delegation; AI augmentation; decision quality

INTRODUCTION

Artificial intelligence has moved from a specialist analytical technology to an increasingly visible participant in managerial work. Machine learning systems can search large information spaces, identify patterns, forecast outcomes, rank alternatives, and generate strategic options at a speed that is difficult for human decision makers to reproduce. Generative AI has widened this role further because managers can now interact with models conversationally, ask for counterarguments, construct scenarios, and obtain rapid evaluations of strategic alternatives. This expansion matters because strategic decisions are typically consequential, uncertain, resource intensive, and difficult to reverse. Their quality depends not only on predictive accuracy but also on framing, interpretation, causal reasoning, political feasibility, ethical legitimacy, and the ability to learn from weak signals. AI therefore creates a genuine strategic opportunity while simultaneously creating a design problem concerning how much judgment should be delegated to the technology and how much should remain embedded in human deliberation (Agrawal et al., 2018; Duan et al., 2019; Jarrahi, 2018; Krakowski et al., 2023; Raisch & Krakowski, 2021). Recent strategy research confirms that AI can contribute meaningfully to strategic evaluation,

but also shows that individual model judgments may be inconsistent or biased, making the architecture of human-AI collaboration central to decision quality (Doshi et al., 2025).

The literature commonly describes two broad modes of human-AI interaction. Delegation occurs when a human assigns an analytical or decision component to an AI system and accepts its output with limited iterative revision, whereas augmentation occurs when machine outputs become inputs into a continuing process of human interpretation, challenge, and refinement. These modes are not simply technical configurations. They redistribute attention, expertise, accountability, and decision authority inside organizations. Delegation can reduce human noise, standardize repeated evaluations, and release managerial capacity, but it can also conceal assumptions and encourage overreliance when the system appears more objective than it is. Augmentation preserves human participation and can combine machine prediction with contextual judgment, yet it may fail when humans anchor on algorithmic recommendations or intervene in ways that reduce rather than improve model performance (Baird & Maruping, 2021; Fügner et al., 2021, 2022; Puranam, 2021; Shrestha et al., 2019). The key question is therefore not whether humans or AI are generally better decision makers. It is how the allocation and sequencing of their respective capabilities influence the quality of strategic choices under different conditions.

Strategic decision quality is itself multidimensional. Accuracy and consistency matter, but strategy also requires fit with organizational capabilities, robustness to uncertainty, recognition of second-order effects, legitimacy among stakeholders, and the capacity to revise assumptions when environments change. Human judgment is vulnerable to bounded rationality, heuristics, bias, noise, fatigue, and political influence, which creates a strong case for computational support (Kahneman, 2011; Kahneman et al., 2021; Simon, 1955; Tversky & Kahneman, 1974). Yet AI systems are trained on historical data and may reproduce data limitations, proxy relationships, distribution shifts, and opaque decision rules. The apparent contrast between imperfect humans and rational machines is therefore misleading. Both parties have distinctive failure modes. Research on algorithm aversion shows that people may reject useful algorithms after observing errors, while algorithm appreciation shows that people can also prefer algorithmic advice over human judgment in some settings (Dietvorst et al., 2015, 2018; Logg et al., 2019). The strategic challenge is to calibrate reliance rather than maximize either human control or machine autonomy.

The problem motivating this review is that empirical evidence on human-AI collaboration is fragmented across strategic management, information systems, organizational behavior, psychology, and computer science. Studies often examine operational forecasting, professional diagnosis, knowledge work, or laboratory judgment and then imply broader managerial relevance. This creates uncertainty about whether findings transfer to strategic decisions, which tend to involve greater ambiguity, irreversibility, political negotiation, and identity stakes. The attached review outline explicitly identifies this fragmentation and the absence of a systematic synthesis that distinguishes delegation from augmentation in strategic rather than merely operational or clinical settings. It also proposes attention to individual, task, and organizational moderators and to the generative AI inflection point. These priorities are retained in the present review because they provide a coherent basis for separating collaboration modes while recognizing that their effects are contingent rather than universal. Recent evidence reinforces this need by showing that top-management use of AI differs according to problem structuredness and situational predictability (Srinivas & Chetan, 2026), while meta-analytic evidence across broader human-AI tasks indicates that combining humans and AI does not automatically outperform the better standalone actor (Vaccaro et al., 2024).

A further gap concerns the level at which collaboration is theorized. Much existing work focuses on whether an individual accepts or rejects algorithmic advice, but strategic decisions are embedded in organizational routines, governance systems, top-management teams, data infrastructures, and accountability arrangements. Organization design research suggests that AI changes who possesses decision rights and how information is processed, while sociotechnical research shows that algorithmic systems become consequential through their interaction with professional practices and institutional controls (Kellogg et al., 2020; Murray et al., 2021; Shrestha et al., 2019; Sturm et al., 2021). The gap is therefore not a simple shortage of studies on AI and decision-making. It is the shortage of integrative evidence explaining how delegation and augmentation produce different decision-quality outcomes once human expertise, task uncertainty, organizational governance, and strategic context are considered together. Generative AI deepens this gap because its conversational and generative

capabilities blur the boundary between tool, adviser, evaluator, and quasi-agent. A review that explicitly compares collaboration modes can therefore clarify both theory and practice.

The aim of this paper is to critically synthesize evidence on how AI delegation and AI augmentation affect the quality of strategic decision-making and to identify the conditions under which each mode is most appropriate. The review has five objectives: to clarify the dominant conceptualizations of delegation and augmentation; to evaluate the effects of AI delegation on strategic decision quality; to evaluate the effects of AI augmentation on strategic decision quality; to identify individual, task, and organizational moderators that explain variation in outcomes; and to develop an integrative framework and future research agenda for generative AI in strategic decision-making. Correspondingly, the review addresses five questions: how are delegation and augmentation conceptualized; what effects does delegation have and under what conditions; what effects does augmentation have and under what conditions; which moderators shape these relationships; and which gaps should guide future research? These objectives extend the structure proposed in the uploaded review outline and maintain its emphasis on strategic decision quality rather than productivity alone. The distinction is important because faster or cheaper decision processes are not necessarily better strategic decisions.

This review contributes in three ways. First, it treats delegation and augmentation as distinct but potentially sequential modes rather than collapsing all AI-assisted work into a single category. Second, it develops a contingency explanation in which decision quality depends on the fit between collaboration mode and the characteristics of the decision environment. Third, it integrates psychological mechanisms such as trust calibration with organizational mechanisms such as governance, decision rights, and learning capacity. The paper is organized as follows. Section 2 critically reviews the conceptual and theoretical literature on strategic decision-making, AI delegation, AI augmentation, decision quality, and relevant moderators. Section 3 explains the PRISMA-guided systematic review methodology, search strategy, eligibility criteria, quality appraisal, and thematic synthesis. Section 4 presents the descriptive and thematic findings from the 25 retained studies and develops an integrative framework. Section 5 discusses theoretical and managerial implications, including the distinctive role of generative AI. Section 6 concludes the review, notes its principal limitations, and identifies future research directions.

LITERATURE REVIEW

Strategic decision-making and bounded rationality

Strategic decision-making concerns choices that commit significant organizational resources and shape long-term direction. Such decisions differ from routine operational choices because objectives can be contested, information is incomplete, outcomes are delayed, and the environment may change before consequences become observable. Simon's bounded rationality perspective remains foundational because it explains why decision makers satisfice rather than optimize when search costs, cognitive limits, and uncertainty constrain choice (Simon, 1955). Behavioral research further demonstrates that human judgment is influenced by heuristics, framing, overconfidence, availability, and noise (Kahneman, 2011; Kahneman et al., 2021; Tversky & Kahneman, 1974). AI appears attractive in this context because computational systems can expand search, process high-dimensional data, and apply consistent rules. However, bounded rationality does not disappear when computation increases. It can shift into data selection, objective specification, model design, prompt construction, and interpretation. AI therefore changes the location of cognitive constraints rather than eliminating them, which makes the distribution of judgment between humans and machines a central strategic design issue.

From an economic perspective, AI lowers the cost of prediction and can therefore alter the value of complementary human judgment (Agrawal et al., 2018). This distinction is particularly useful for strategy because prediction and judgment are not identical. A model may estimate demand, competitive response, or the probability of success, while executives must still decide which outcomes matter, how risks should be traded off, and whether the organization can implement a chosen course. Generative AI complicates the distinction by producing narratives and recommendations that resemble judgment, even though their outputs remain grounded in learned statistical patterns. Doshi et al. (2025) show that generative AI can provide strategically useful evaluations, especially when multiple assessments are aggregated, but individual evaluations can be inconsistent and biased. Felin and Holweg (2024) similarly emphasize the continuing importance of human causal reasoning

and theory. Strategic decision quality should therefore be assessed not only by prediction but by the relationship between prediction, explanation, context, and accountable choice.

AI delegation and the redistribution of decision authority

Delegation can be understood as assigning a task, analysis, or decision component to an agentic information system that performs work on behalf of a human actor. Baird and Maruping (2021) frame delegation as a relationship involving goals, authority, and monitoring between human users and agentic information systems. In strategic contexts, delegation may include market screening, ranking acquisition targets, scoring investment options, generating forecasts, or selecting among alternatives when criteria are explicit. Its attraction is straightforward. Algorithms can process more alternatives, apply criteria consistently, and reduce some forms of interpersonal and temporal variability. Kim et al. (2024) highlight the importance of decision authority in determining returns to algorithms, while Shrestha et al. (2019) show that AI can reshape organizational decision structures. Delegation therefore has the potential to improve quality when a decision component is well specified and the model's objective aligns with strategic goals. The risk is that authority can be transferred more quickly than accountability, understanding, or the ability to contest the output.

Delegation is also psychologically attractive because it can reduce cognitive burden and create an appearance of neutrality. Yet this appearance may encourage automation bias or responsibility displacement. Keding and Meissner (2021) show that AI-based advisory systems can shape managerial choice behavior and create overreliance in R&D investment decisions. Fügener et al. (2022) demonstrate that productive delegation requires managing cognitive challenges rather than assuming that humans will naturally know when to rely on AI. Delegation can also weaken learning if managers stop practicing the interpretive skills required to recognize model failure. Lebovitz et al. (2021) warn that the ground truth used to train and evaluate AI may itself reflect contested expert knowledge, while Teodorescu et al. (2021) show that fairness failures cannot be understood solely as machine problems because human-machine augmentation shapes outcomes. These findings imply that delegation is most defensible when task boundaries, evaluation criteria, feedback loops, and escalation rules are explicit.

AI augmentation and complementary intelligence

Augmentation treats AI as a collaborator that extends rather than replaces human cognition. Jarrahi (2018) describes human-AI symbiosis as a way to combine computational analysis with human intuition and contextual understanding. Raisch and Krakowski (2021) develop the automation-augmentation paradox, arguing that organizations face interdependent pressures to automate tasks while also enhancing human capabilities. Augmentation can occur through specialization, where AI handles prediction and humans handle contextual judgment, or through iterative interaction in which each party repeatedly modifies the other's contribution. Shrestha et al. (2021) argue that deep learning can augment organizational decision-making but requires attention to explainability, accountability, and task allocation. Puranam (2021) goes further by framing collaborative decision-making as an organization design problem. The central promise is complementarity: decision quality can improve when human and machine errors are imperfectly correlated and when each party contributes information or reasoning unavailable to the other.

Complementarity, however, should not be presumed. Fügener et al. (2021) show that humans-in-the-loop can produce both merits and pitfalls, including situations in which interaction degrades performance. Bansal et al. (2021) similarly demonstrate that the most accurate standalone AI is not necessarily the best teammate because collaboration quality depends on whether human users can understand and use the model's strengths. Vaccaro et al. (2024), in a systematic review and meta-analysis, find that human-AI combinations do not consistently outperform the better of humans or AI alone, even though they often outperform humans alone. This is a critical corrective to optimistic augmentation narratives. The relevant comparison is not whether AI helps humans, but whether a designed human-AI system produces better outcomes than credible alternatives. Augmentation therefore requires complementary expertise, calibrated trust, useful explanations, and interaction protocols that allow disagreement to generate information rather than confusion.

Algorithm aversion, appreciation, trust, and expertise

Human responses to AI are neither uniformly skeptical nor uniformly deferential. Dietvorst et al. (2015) demonstrate algorithm aversion, where people abandon algorithms after seeing them make errors even when the algorithms remain superior to human forecasters. Their later work shows that allowing users even slight control over an algorithm can increase willingness to use it (Dietvorst et al., 2018). In contrast, Logg et al. (2019) document algorithm appreciation, showing that people sometimes weight algorithmic advice more heavily than human advice. These apparently contradictory patterns are better interpreted as evidence that reliance is context dependent. Expertise, perceived control, task familiarity, outcome accountability, and the visibility of errors all influence trust. For strategic decisions, the ideal is calibrated trust, meaning reliance that tracks actual system competence. Under-trust wastes useful analytical capacity, while over-trust allows model limitations to enter strategic commitments without adequate challenge. Collaboration design should therefore make confidence, uncertainty, and conditions of model validity visible to decision makers.

Domain expertise has a similarly complex role. Experts possess tacit knowledge and causal models that can help identify implausible outputs, but expertise can also intensify resistance when AI threatens professional identity or contradicts established routines. Allen and Choudhury (2022) describe countervailing forces of ability and aversion in algorithm-augmented work, while Wang et al. (2023) show that the effects of teaming with AI vary with worker experience. Lebovitz et al. (2022) find that professionals confronted with opaque AI engage selectively, depending on how they can reconcile system outputs with their own critical judgments. In strategy, senior executives may be especially vulnerable to both sides of this tension because they possess contextual knowledge but also strong identities as decision authorities. AI literacy is therefore not simply technical training. It includes understanding model uncertainty, recognizing hallucination and distribution shift, knowing when independent verification is needed, and retaining the confidence to reject outputs when strategic context provides contrary evidence.

Task characteristics and the delegation-augmentation boundary

Task structure is one of the strongest theoretical candidates for explaining when delegation or augmentation should dominate. Structured tasks have clear inputs, measurable criteria, repeated observations, and relatively stable relationships between variables. These conditions support delegation because performance can be evaluated and corrected. Ill-structured strategic problems involve ambiguous goals, novel competitors, political trade-offs, sparse data, and uncertain causal relationships. In such settings, historical pattern recognition may be informative but insufficient. Recent executive-level evidence suggests that structured and predictable decision components are more amenable to AI delegation, whereas ill-structured or uncertain strategic problems call for iterative human-AI co-deliberation (Srinivas & Chetan, 2026). This distinction is consistent with broader arguments about bounded rationality and organization design. It also clarifies why debates that ask whether AI should replace managers are poorly specified. The appropriate unit of analysis is often the decision component, not the entire strategic decision.

Decision reversibility and consequence severity further shape the collaboration boundary. When a decision is easily reversible, organizations can tolerate experimentation and learn from AI errors. When decisions involve acquisitions, market exits, major capital investments, workforce restructuring, or public commitments, errors can be costly and difficult to undo. Strategic decisions also contain social and political dimensions that may not be represented in available data. Consequently, a model that performs well on predictive criteria can still recommend an option that is infeasible, illegitimate, or inconsistent with organizational identity. The value of augmentation increases when decision makers must integrate quantitative predictions with tacit, ethical, political, and implementation knowledge. Conversely, delegation can remain valuable for bounded subproblems within the same decision, such as competitor scanning or sensitivity analysis. This suggests that hybrid, staged architectures may be more realistic than a binary choice between full automation and full human control.

Organizational governance, learning, and accountability

AI collaboration is embedded in organizational structures. Li et al. (2021) emphasize the role of CIOs and boards in setting strategic directions for AI, while Benbya et al. (2021) show that AI raises broad questions for information systems research concerning organizational use and governance. Governance matters because

delegation requires clarity about who authorizes the model, who validates data, who can override outputs, and who remains accountable for consequences. Kellogg et al. (2020) demonstrate that algorithms can become a contested terrain of control, affecting autonomy and power rather than merely efficiency. Strategic AI governance should therefore include decision-rights design, documentation, auditability, escalation mechanisms, and periodic review of model performance. These controls are not peripheral compliance activities. They influence decision quality by determining whether weak signals and dissenting judgments can challenge an apparently authoritative model output.

Organizational learning capacity is equally important. Sturm et al. (2021) argue that effective organizational learning requires coordination between human and machine learning. If AI systems update without corresponding changes in managerial understanding, organizations can become dependent on outputs they cannot interpret. Conversely, if managers learn but models remain static, technological support may become progressively misaligned with changing environments. Brynjolfsson et al. (2021) show more broadly that productivity benefits from general-purpose technologies often require complementary intangible investments. Applied to strategic AI, these investments include data governance, training, redesigned workflows, evaluation routines, and mechanisms for sharing lessons from model success and failure. Augmentation is therefore an organizational capability rather than a simple software feature. Firms that lack these complements may obtain little value from sophisticated AI or may use it in ways that create false confidence.

Generative AI and the changing nature of strategic collaboration

Generative AI changes the human-AI relationship because it can participate in language-intensive activities that were previously difficult to automate, including scenario generation, synthesis, argument construction, and evaluation. Noy and Zhang (2023) demonstrate productivity effects of generative AI in professional writing tasks, while Dell'Acqua et al. (2023) show that performance benefits depend on whether tasks fall inside or outside the technology's uneven capability frontier. Lee and Chung (2024) add evidence on creativity, illustrating that generative systems can influence idea production as well as analysis. For strategic decision-making, these capabilities allow AI to operate earlier in the process, before alternatives are fully specified. This is consequential because problem framing and option generation shape the eventual choice set. A system that appears to augment decision-making may quietly structure the decision by determining which assumptions, competitors, risks, or alternatives receive attention.

The generative AI literature also sharpens concerns about reliability. Doshi et al. (2025) find that individual generative AI evaluations of strategic alternatives can be inconsistent and biased, although aggregation across models, prompts, or roles improves correspondence with expert rankings. Simkute et al. (2024) highlight ironies in human-generative AI interaction, including circumstances in which expected productivity gains can be undermined by interaction costs. These findings suggest that generative AI should not be treated as a single stable adviser. Prompt wording, model version, role instruction, and sampling variation can change outputs. Strategic use therefore benefits from triangulation, independent prompts, model ensembles, and explicit human verification. The emerging research gap is no longer simply whether AI can support strategy. It concerns how organizations should design repeatable, accountable collaboration processes when the AI itself is generative, probabilistic, and capable of producing persuasive but unsupported reasoning.

Critical synthesis and research gap

The synthesis draws on four theoretical lenses that recur across the retained literature and that structure the interpretation offered in this review. Table 1 summarizes these lenses, their principal sources, and the role each plays in explaining delegation, augmentation, and decision quality.

Table 1: Theoretical lenses guiding the review

Theoretical lens	Principal sources	Role in this review
Bounded rationality	Simon (1955); Kahneman (2011); Kahneman, Sibony, and Sunstein (2021); Tversky and Kahneman (1974)	Explains why human judgment satisfices under cognitive and informational limits, and why AI relocates rather than eliminates these constraints

Algorithm aversion and appreciation	Dietvorst, Simmons, and Massey (2015, 2018); Logg, Minson, and Moore (2019)	Explains inconsistent human reliance on AI advice and motivates the review's emphasis on calibrated trust
Organization design	Puranam (2021); Shrestha, Ben-Menahem, and von Krogh (2019); Baird and Maruping (2021)	Frames human-AI collaboration as a problem of task allocation, decision rights, and coordination rather than simple tool adoption
Automation-augmentation paradox	Raisch and Krakowski (2021)	Explains why organizations face simultaneous pressure to automate tasks and to enhance human capability, motivating the review's contingency framework

The literature establishes that both delegation and augmentation can improve decision processes, but it does not support a universal superiority claim for either mode. Delegation is strongest where prediction tasks are structured, measurable, and stable, while augmentation becomes more valuable as ambiguity, novelty, irreversibility, and contextual dependence increase. Yet even this contingency view remains incomplete because individual trust, expertise, and organizational governance can change the effect of the same technical configuration. A highly accurate system can be a poor teammate, and a nominal human-in-the-loop can add little protection when users routinely accept outputs without independent reasoning (Bansal et al., 2021; Fügener et al., 2021; Vaccaro et al., 2024). The central gap is therefore the absence of an integrated strategic-level synthesis that simultaneously distinguishes delegation from augmentation, evaluates their effects on multidimensional decision quality, and explains moderation across individual, task, and organizational levels. The present systematic review addresses this gap by synthesizing the retained evidence thematically and by proposing a framework in which collaboration mode is matched to decision conditions rather than selected as a general organizational philosophy.

METHODOLOGY

A systematic literature review was selected because the relevant evidence is dispersed across several disciplines and uses heterogeneous concepts and methods. A systematic approach improves transparency, reproducibility, and critical comparison relative to an unsystematic narrative review (Snyder, 2019; Tranfield et al., 2003). Reporting was guided by PRISMA 2020 because it provides a recognized structure for documenting identification, screening, eligibility, and inclusion decisions (Page et al., 2021). The review was interpretive rather than meta-analytic because the retained studies differed substantially in research design, collaboration mode, task context, and measures of decision quality. The protocol comprised six documented elements, addressed in turn in the subsections below: search strategy and information sources, eligibility criteria, screening and PRISMA flow, data extraction and coding, quality appraisal, and thematic synthesis.

Search strategy and information sources

Searches combined AI terms with strategic decision and collaboration terms. The core logic was: ("artificial intelligence" OR "AI" OR "algorithm*" OR "machine learning" OR "large language model*" OR "generative AI") AND ("strategic decision*" OR "managerial decision*" OR "decision-making") AND ("delegation" OR "augmentation" OR "human-AI collaboration" OR "decision quality"). Google Scholar, Scopus, Web of Science, and ResearchGate were used in the identification stage shown in Figure 1. Multiple sources were justified because the field spans management, information systems, psychology, and computing, making single-database coverage inadequate. Searches were conducted in September 2026 and were restricted to publications from January 2015 to September 2026, a period spanning the emergence of enterprise-scale predictive AI in managerial decision support through to the most recent generative AI developments.

Table 2: Search strategy and rationale

Element	Operationalization	Justification
Population/context	Managers, executives, organizations, strategic or managerial decisions	Keeps the review focused on consequential organizational decision settings.
AI concepts	AI, algorithm*, machine learning, LLM*, generative AI	Captures both established predictive AI and newer generative systems.
Collaboration concepts	delegation, augmentation, human-AI collaboration, decision support	Separates authority transfer from collaborative judgment.
Outcome concepts	decision quality, accuracy, consistency, strategic fit, judgment	Reflects multidimensional strategic decision quality.
Sources	Google Scholar, Scopus, Web of Science, ResearchGate	Broadens interdisciplinary coverage and reduces database-specific omission.

Eligibility criteria

Studies were eligible when they examined AI in relation to human or organizational decision-making and provided evidence or theory relevant to delegation, augmentation, or decision quality. Priority was given to peer-reviewed empirical and conceptual work in management, information systems, organizational behavior, and closely related fields. Studies focused exclusively on low-level technical model performance without a human or organizational decision component were excluded. Clinical and operational studies were retained only when they offered mechanisms clearly transferable to strategic collaboration, such as opacity, trust, or expert engagement. English-language publication was used to support consistent interpretation and coding.

Table 3: Inclusion and exclusion criteria

Criterion	Included	Excluded
Topical relevance	Human-AI collaboration, delegation, augmentation, strategic or managerial decision quality	Purely technical AI optimization without a human decision component
Context	Strategic, managerial, organizational, or theoretically transferable professional judgment	Contexts with no plausible relevance to organizational decision-making
Evidence	Peer-reviewed empirical, theoretical, and high-value systematic evidence	Editorial commentary without analytical contribution
Language	English	Non-English where reliable full-text assessment was not possible
Outcome	Decision quality or mechanisms affecting quality	Studies concerned only with adoption intention without decision implications

Screening and PRISMA flow

The identification stage produced 4,087 records. After 1,027 duplicates were removed, 3,060 titles were screened. A total of 2,420 records were excluded at this stage, leaving 640 full-text articles sought for retrieval. Of these, 112 could not be retrieved, so 528 full texts were assessed for eligibility. A further 503 records were

excluded because they did not sufficiently address AI delegation or augmentation, decision quality, strategic or organizational context, empirical or peer-reviewed relevance, or methodological detail. The final synthesis retained 25 studies. Figure 1 reports this process and provides an auditable link between the search universe and the evidence base.

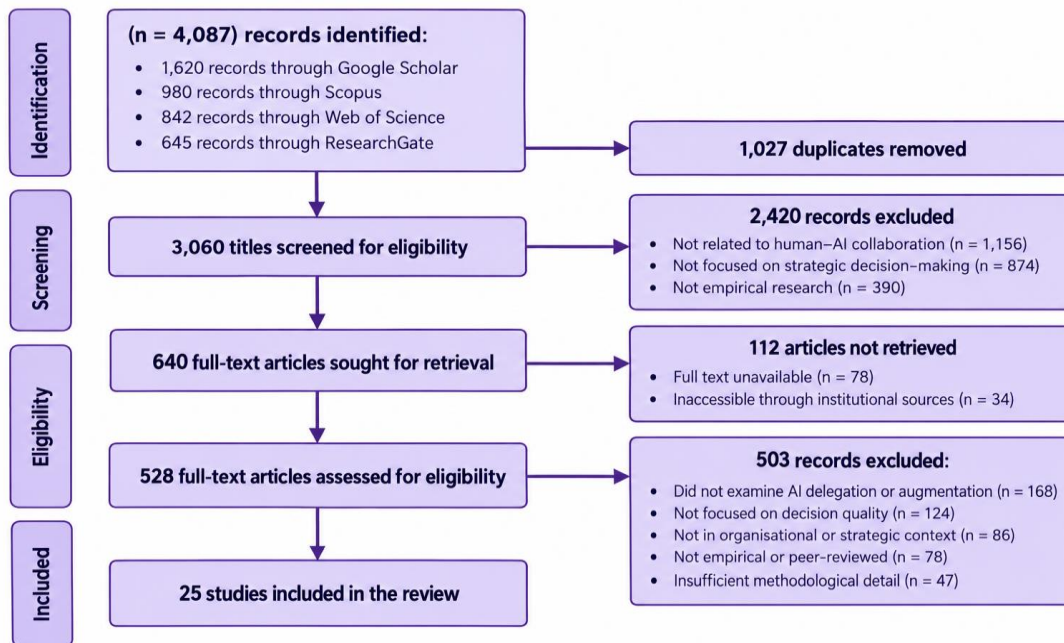


Figure 1. PRISMA flow diagram for study identification, screening, eligibility, and inclusion.

Data extraction and coding

A structured extraction form was used to record author and year, disciplinary setting, theoretical lens, research design, AI engagement mode, decision context, decision-quality measure, moderators, and principal findings. This structure was chosen because the review questions require comparison across both outcomes and mechanisms. Delegation and augmentation were coded as analytical categories rather than assuming that authors used identical terminology. Where studies described mixed arrangements, coding reflected the dominant mode and the sequencing of human and AI contributions. This approach reduced the risk that differences in labels would obscure substantively similar collaboration designs.

Table 4: Data extraction framework

Field	Coding focus
Bibliographic profile	Author, year, outlet, discipline
Decision context	Strategic, managerial, professional, experimental
AI engagement mode	Delegation, augmentation, hybrid or staged
Human contribution	Framing, contextual judgment, verification, override, final authority
AI contribution	Prediction, search, classification, generation, evaluation, recommendation
Decision-quality outcome	Accuracy, consistency, strategic fit, creativity, robustness, accountability
Moderators	Expertise, trust, task structure, uncertainty, governance, AI literacy
Key conclusion	Direction, boundary conditions, risks, implications

Figure 2 depicts this extraction and coding schema as a single flow, showing how each field builds on the one before it, from bibliographic identification through to the key conclusion recorded for each study.

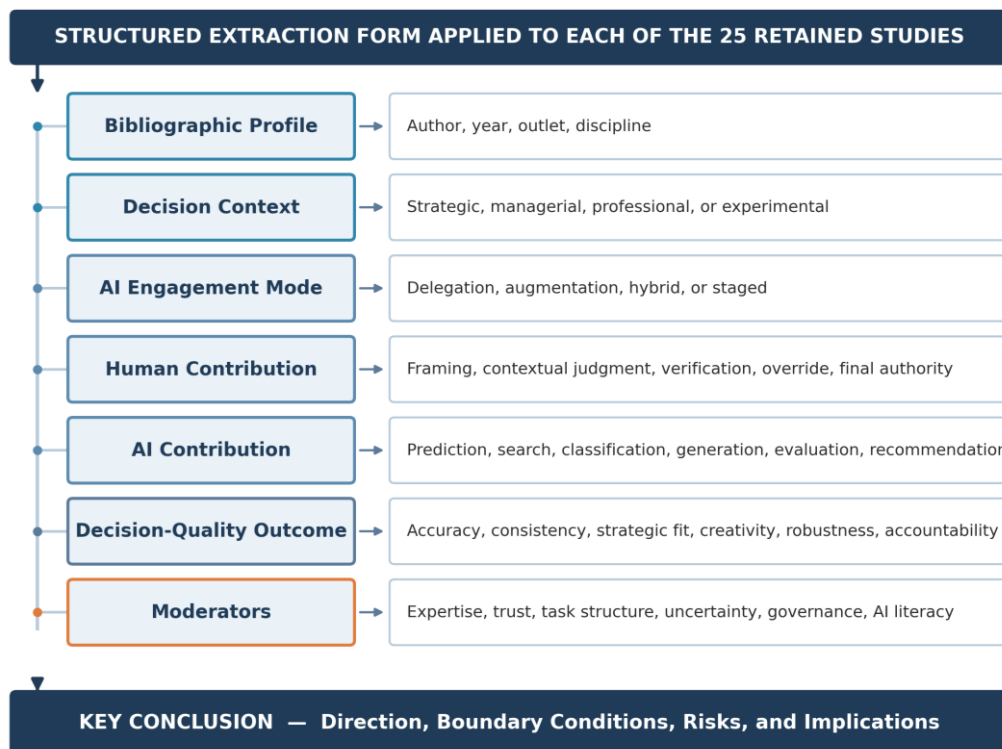


Figure 2. Data extraction and coding schema applied to the 25 retained studies.

Quality appraisal

Quality appraisal considered clarity of research question, appropriateness of design, transparency of sampling or data sources, validity of outcome measures, analytical coherence, and the extent to which conclusions followed from evidence. A single numerical score was not used because the corpus included experiments, field studies, conceptual papers, and review evidence that cannot be judged fairly with one metric. Instead, quality appraisal informed interpretive weight. Strong causal claims were reserved for designs that supported them, while conceptual studies were used primarily to explain mechanisms and boundary conditions. This decision protects the synthesis from treating all published claims as equally evidential.

Thematic synthesis

The retained evidence was synthesized through thematic analysis. Initial codes captured recurrent mechanisms such as consistency gains, overreliance, complementarity, trust calibration, task ambiguity, and governance. Codes were then grouped into higher-order themes that directly answered the review questions. Thematic synthesis was preferred because statistical pooling would have obscured differences in constructs and contexts. The procedure followed the logic of iterative familiarization, coding, theme development, review, and definition associated with thematic analysis (Braun & Clarke, 2006), while maintaining a critical narrative emphasis appropriate to management reviews (Snyder, 2019).

FINDINGS AND DISCUSSION

The 25 retained studies show a field moving from a simple comparison of human versus algorithmic performance toward a more relational question about how decision work should be divided and coordinated. Across the corpus, AI contributes most consistently through scale, speed, prediction, search, and reduction of some forms of judgmental variance. Humans contribute most consistently through problem framing, contextual interpretation, causal reasoning, ethical evaluation, stakeholder awareness, and accountability. The thematic analysis produced five interrelated themes: conditional gains from delegation; complementarity and error

correction through augmentation; trust calibration and expertise; task structure, uncertainty, and reversibility; and organizational governance and learning. These themes are summarized in Table 5 and developed below. The synthesis is deliberately critical because positive average effects can conceal collaboration failures, while negative findings can reflect poor task allocation rather than intrinsic limits of AI.

Table 5: Thematic synthesis of the retained literature

Theme	Dominant evidence pattern	Critical interpretation
Conditional value of delegation	Speed, consistency, scalable evaluation, reduced noise	Benefits depend on structured tasks, reliable data, and monitorable outputs.
Augmentation and complementarity	Human context plus machine prediction can improve judgment	Human involvement adds value only when it contributes independent information or effective correction.
Trust and expertise	Aversion and overreliance coexist	Decision quality requires calibrated reliance, not maximum trust.
Task structure and uncertainty	Delegation fits structured tasks; augmentation fits ambiguous and irreversible problems	The relevant unit is often the decision component rather than the whole strategic decision.
Governance and learning	Decision rights, auditability, AI literacy, and feedback shape outcomes	Human-in-the-loop is insufficient without real contestability and accountability.

Figure 3 maps these five themes and their dominant evidence patterns before each is developed in turn below.

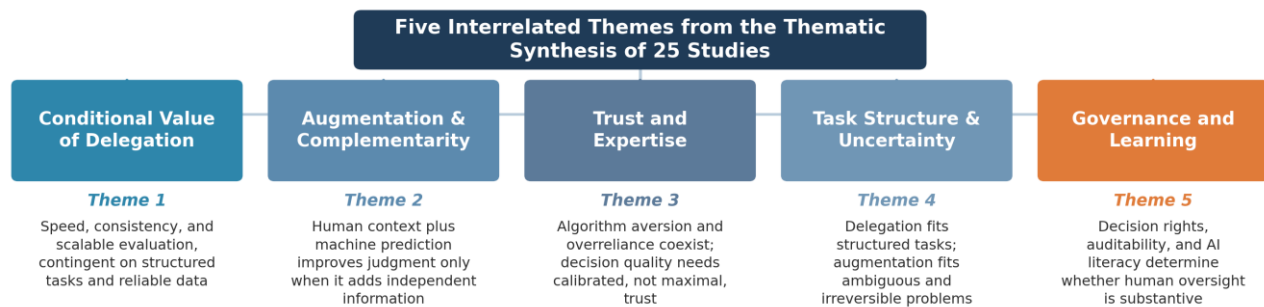


Figure 3. Thematic map of the five themes emerging from the retained literature.

Theme 1: Conditional value of AI delegation

The first theme is that delegation can improve decision quality when the delegated component is structured and its success criteria are observable. Algorithms can evaluate more alternatives than managers, apply the same criteria repeatedly, and reduce noise caused by fatigue or inconsistent human attention. This advantage aligns with the economics of prediction and with behavioral evidence showing substantial variability in human judgment (Agrawal et al., 2018; Kahneman et al., 2021). Delegation is therefore particularly defensible for bounded analytical tasks such as screening, scoring, forecasting, and sensitivity analysis. Kim et al. (2024) show that the allocation of decision authority is consequential for the returns organizations obtain from algorithms, while Baird and Maruping (2021) clarify that delegation involves more than mere tool use because authority and monitoring relationships change. The critical qualification is that strategic value depends on whether the delegated task is genuinely specifiable. A model can be internally consistent while optimizing an incomplete or inappropriate objective.

The negative side of delegation appears when managerial actors confuse consistency with correctness or treat computational output as neutral. Keding and Meissner (2021) demonstrate that AI-based advice can create managerial overreliance, and Fügener et al. (2022) show that productive delegation involves cognitive challenges that must be actively managed. Delegation can also narrow attention if managers accept the model's representation of the problem without questioning excluded variables or changing environmental conditions. This is especially dangerous in strategy because the most important information may be novel, weakly structured, politically sensitive, or absent from historical data. Lebovitz et al. (2021) further challenge the assumption that training labels represent uncontested truth. The review therefore finds that delegation improves strategic decision quality only when organizations maintain verification, monitor distribution shifts, preserve escalation pathways, and recognize that model outputs remain conditional on data and objectives.

Theme 2: Augmentation, complementarity, and error correction

The second theme is that augmentation can create value through cognitive division of labor. AI can perform high-volume search and prediction while humans interpret meaning, incorporate tacit knowledge, assess causal plausibility, and consider stakeholder consequences. This pattern is consistent with Jarrahi's (2018) account of human-AI symbiosis and Puranam's (2021) treatment of collaboration as an organization design problem. Choudhury et al. (2020) show that machine learning and human capital can be complementary in ways that mitigate bias, while Shrestha et al. (2021) identify principles and challenges for augmenting organizational decisions with deep learning. The strongest argument for augmentation is not that human judgment is always superior on context, but that humans and AI may possess different information and make different errors. When those errors are imperfectly correlated, combining them can improve robustness.

However, the evidence rejects a simplistic assumption that adding a human to AI automatically creates superior performance. Fügener et al. (2021) show that human-in-the-loop arrangements can produce pitfalls, and Bansal et al. (2021) demonstrate that optimizing an AI for standalone accuracy can make it a poorer teammate. Vaccaro et al. (2024) provide especially important aggregate evidence: human-AI combinations often improve on humans alone but do not consistently beat the better standalone performer. This means augmentation must be designed to create genuine complementarity. If humans merely confirm AI outputs, the arrangement is functionally delegated even if a formal approval step remains. If humans override accurate outputs because of aversion or ego, collaboration can also degrade quality. Effective augmentation therefore requires independent human reasoning, visibility of uncertainty, meaningful opportunities to contest outputs, and clear rules for resolving disagreement.

Theme 3: Trust calibration, expertise, and cognitive responses

The third theme concerns trust calibration. Algorithm aversion and algorithm appreciation coexist because reliance depends on experience, perceived competence, error visibility, and control. Dietvorst et al. (2015) show that observing algorithmic error can lead users to reject systems that remain objectively useful, while Dietvorst et al. (2018) show that limited user control can restore acceptance. Logg et al. (2019) demonstrate the opposite tendency, where people may prefer algorithmic judgment. In strategic settings, both underreliance and overreliance are costly. Underreliance wastes computational insight and can preserve human bias. Overreliance allows persuasive or precise outputs to bypass scrutiny. The appropriate target is calibrated trust, where reliance changes with demonstrated competence, uncertainty, and task conditions. This requires managers to understand not only what a model recommends but where it is likely to fail.

Expertise moderates this relationship in non-linear ways. Experienced decision makers can identify implausible outputs and supply context that data do not capture, but they may also resist AI when it threatens professional identity or established authority. Allen and Choudhury (2022) describe ability and aversion as countervailing forces, and Wang et al. (2023) show that experience changes the consequences of AI teaming. Lebovitz et al. (2022) illustrate how professionals engage with opaque AI selectively rather than simply accepting or rejecting it. For executives, this implies that AI literacy must include critical engagement skills. Training should develop the ability to test assumptions, request alternative scenarios, compare outputs across prompts or models, and recognize when strategic context invalidates historical patterns. Expertise should therefore be treated as a resource for structured contestation rather than as a reason either to ignore AI or to defer to it.

Theme 4: Task structure, uncertainty, and reversibility

The fourth theme provides the clearest contingency logic. Delegation is most defensible when problems are structured, criteria are explicit, and the environment is sufficiently predictable for historical relationships to remain informative. Augmentation becomes more important as tasks become ambiguous, novel, politically contested, or difficult to reverse. Srinivas and Chetan (2026) provide recent top-management evidence consistent with this distinction, finding that executives differentiate between high-confidence delegation of analytical work and iterative co-deliberation for ill-structured strategic problems. The distinction also explains why operational success with AI cannot be directly generalized to strategy. A demand forecast may have clear error metrics and frequent feedback, whereas a market-entry decision may be evaluated only after years and may be confounded by implementation, competitor response, and macroeconomic change. Strategic uncertainty therefore increases the value of human interpretation and governance.

Reversibility strengthens this pattern because it changes the cost of model error. When decisions can be tested cheaply, organizations can delegate more aggressively and learn from outcomes. When choices involve major capital commitments, acquisitions, reputational exposure, or workforce consequences, the value of contestability rises. This does not mean AI should be excluded from high-stakes decisions. Instead, its role should shift toward scenario exploration, evidence synthesis, counterfactual generation, and structured challenge. Doshi et al. (2025) illustrate this logic by showing that generative AI can evaluate strategic alternatives, particularly when multiple evaluations are aggregated, but single assessments can be biased or inconsistent. The finding supports a staged architecture in which AI expands and evaluates the option space while humans retain responsibility for contextual interpretation and commitment.

Theme 5: Governance, accountability, and organizational learning

The fifth theme is that collaboration quality is an organizational property as much as an individual one. Decision rights determine who can authorize AI use, who can override outputs, and who remains responsible for consequences. Shrestha et al. (2019) show that AI can reshape organizational decision-making structures, while Li et al. (2021) emphasize the strategic role of CIOs and boards. Kellogg et al. (2020) demonstrate that algorithms can alter control and professional autonomy, which means governance affects both acceptance and the quality of challenge. A formal human approval step is insufficient when incentives, workload, or organizational culture encourage automatic acceptance. Effective governance requires traceability, documentation of model limitations, escalation rules, periodic validation, and explicit ownership of final decisions.

Organizational learning determines whether collaboration improves over time. Sturm et al. (2021) argue that human and machine learning must be coordinated, and Brynjolfsson et al. (2021) show that returns from general-purpose technologies depend on complementary intangible investments. In the AI context, those investments include data quality, training, process redesign, experimentation, and mechanisms for sharing lessons. A firm that deploys advanced models without developing these complements may accelerate poor decisions rather than improve them. Governance should therefore be adaptive. It should capture near misses, document overrides, compare predicted and realized outcomes where possible, and revise collaboration protocols as model capabilities and strategic environments change. This is particularly important for generative AI because models, prompts, and platform capabilities can change rapidly.

Generative AI as an inflection point

Generative AI expands the range of strategic activities in which collaboration is possible. Unlike traditional predictive models that often operate on predefined inputs and outputs, large language models can help frame problems, generate options, critique assumptions, summarize external evidence, simulate stakeholder perspectives, and draft strategic narratives. Noy and Zhang (2023) demonstrate that generative AI can improve productivity in knowledge work, while Dell'Acqua et al. (2023) show that benefits depend on whether tasks fall within the technology's capability frontier. This creates a strategic opportunity because managers can use AI not only after alternatives are defined but during the construction of the decision itself. It also creates a new risk: the

model can influence which alternatives are imagined and which assumptions appear salient before formal evaluation begins.

The evidence therefore supports using generative AI as a structured thought partner rather than an unquestioned strategic authority. Doshi et al. (2025) show that aggregation across AI evaluations can reduce inconsistency, suggesting practical techniques such as multiple prompts, independent model roles, and ensemble judgments. Simkute et al. (2024) caution that generative AI interaction can also create productivity losses, which reminds organizations that conversational ease is not equivalent to decision quality. The most defensible design is iterative and adversarial in a constructive sense: managers should ask AI to produce alternatives, critique its own recommendations, identify missing evidence, and generate counterarguments, while humans independently evaluate strategic fit and consequences. This approach treats generative AI as a source of cognitive variety without confusing linguistic fluency with causal understanding.

Integrative framework

Figure 4 integrates the review findings. AI engagement mode influences decision quality through collaboration mechanisms, but the relationship is moderated by individual, task, and organizational conditions. Delegation is expected to perform best when analytical tasks are structured, model validity is high, and outcomes can be monitored. Augmentation is expected to perform best when strategic problems require contextual interpretation, causal reasoning, ethical judgment, or negotiation. Hybrid designs are appropriate when a strategic decision contains both kinds of components. AI engagement mode is, in effect, a decision-authority choice: how much of the analytical and evaluative work is transferred to the system versus retained by the human decision-maker. Individual expertise, trust calibration, and the explainability of AI outputs determine whether recommendations are challenged appropriately; task ambiguity and reversibility determine the cost of mistaken reliance; and organizational readiness, expressed through governance, decision rights, and AI literacy, determines whether human oversight is substantive rather than nominal. These moderators shape the collaboration mechanisms through which AI capabilities such as prediction, search, and generation are accepted, challenged, or overridden. Decision quality is therefore conceptualized broadly to include accuracy, consistency, strategic fit, novelty, accountability, and robustness rather than speed alone.

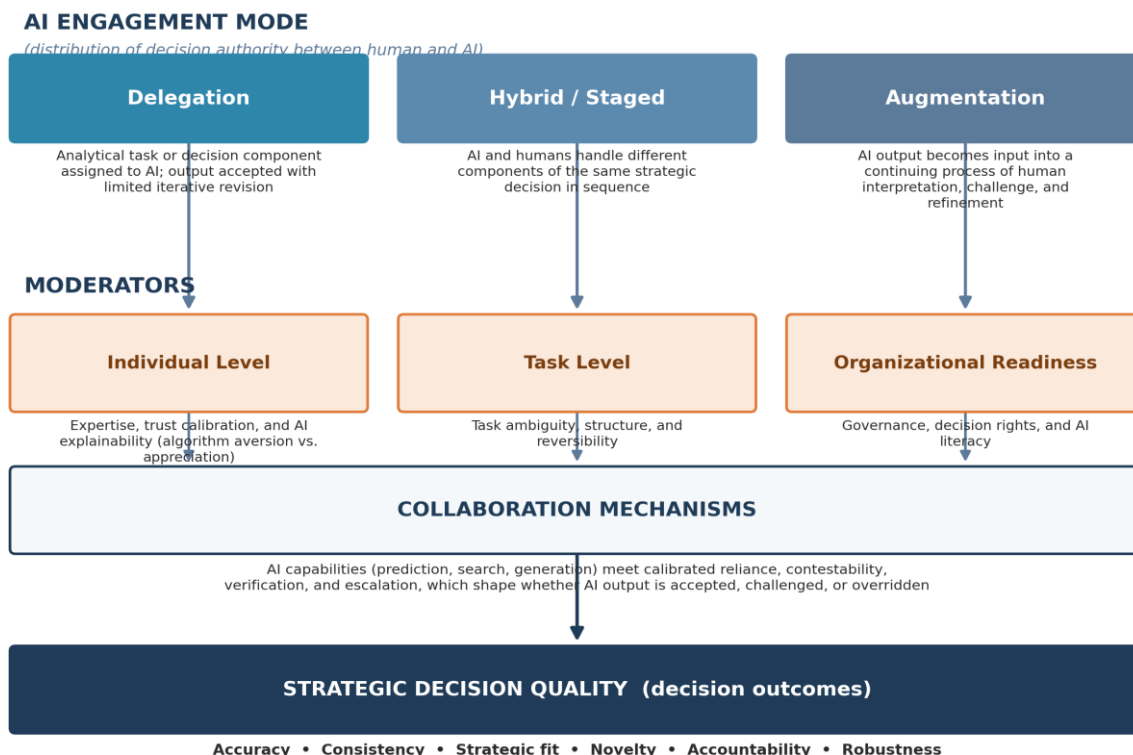


Figure 4. Integrative framework linking AI engagement mode, moderators, collaboration mechanisms, and strategic decision quality.

Comparative synthesis: when delegation or augmentation is preferable

Table 6: Contingent choice between delegation and augmentation

Decision condition	Preferred emphasis	Reason
Structured, repeated, measurable	Delegation	AI can exploit scale and consistency with observable feedback.
Structured but changing	Hybrid	AI provides analysis while humans monitor drift and update assumptions.
Ambiguous, novel, weakly structured	Augmentation	Human framing and contextual interpretation become central.
High irreversibility or stakeholder consequence	Augmentation with strong governance	Contestability and accountability outweigh speed.
Time pressure with reliable model	Delegation of bounded subtask	Releases scarce managerial attention without transferring the whole decision.
Generative exploration	Augmentation and ensemble checking	AI broadens options but outputs require triangulation and causal scrutiny.

Figure 5 renders this contingent architecture as a single diagram linking each decision condition to its preferred collaboration emphasis and underlying rationale.

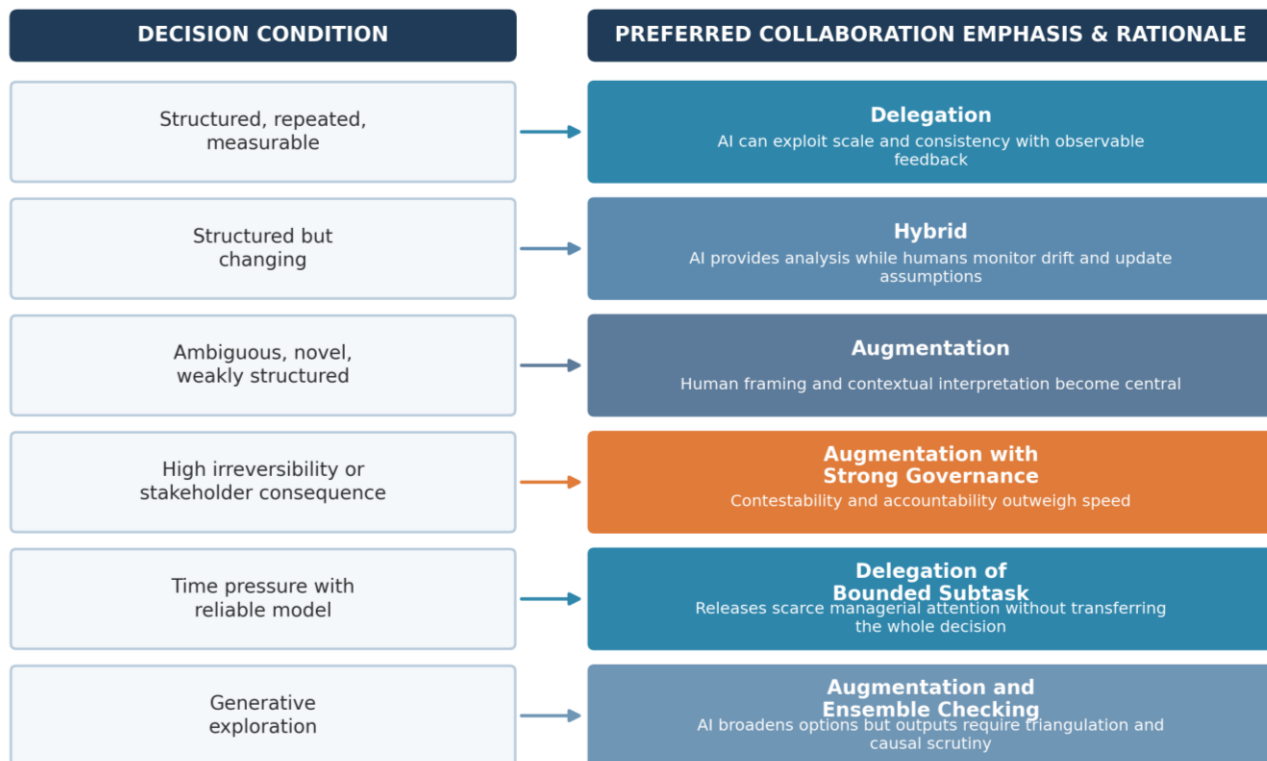


Figure 5. Contingent decision architecture linking decision conditions to the preferred collaboration emphasis.

The comparative synthesis rejects a single optimal collaboration model. Delegation and augmentation are better understood as configurable elements of a decision architecture. A firm can delegate data gathering and forecasting, augment scenario interpretation, and retain human authority over final commitment. This staged approach reconciles the automation-augmentation paradox because automation at one stage can increase the

value of human judgment at another (Raisch & Krakowski, 2021). It also responds to evidence that standalone accuracy does not determine teamwork quality (Bansal et al., 2021). Managers should therefore map the strategic decision process into components, identify which components have reliable data and measurable objectives, and assign AI authority accordingly. Where uncertainty is high, the system should be designed to expose assumptions and alternatives rather than converge prematurely on a recommendation. The practical principle is fit, not maximal automation.

Theoretical implications

The review contributes to bounded rationality theory by showing that AI expands search and prediction while relocating rather than eliminating cognitive constraints. Humans remain responsible for objective selection, causal interpretation, and commitment, while AI introduces new limits related to data representation and model validity. The review also extends algorithm aversion and appreciation research by placing trust within a multilevel organizational context. Reliance is shaped not only by individual psychology but by governance, professional identity, decision rights, and the reversibility of strategic choices. Finally, the findings support organization design perspectives that treat human-AI collaboration as a problem of task allocation and coordination. The theoretical implication is that decision quality emerges from the configuration of actors, information, and authority rather than from the intrinsic superiority of either humans or AI (Murray et al., 2021; Puranam, 2021; Shrestha et al., 2019).

Managerial implications

Executives should begin with decision decomposition rather than technology selection. Strategic processes should be separated into framing, information gathering, prediction, option generation, evaluation, negotiation, and commitment. AI can then be assigned to components where its comparative strengths are clear. Organizations should also establish calibration routines, including independent checks, confidence reporting, model comparison, and post-decision review. For generative AI, managers should avoid relying on a single prompt or a single model response for consequential judgments. Aggregation and adversarial prompting can reduce some instability, although they do not eliminate bias (Doshi et al., 2025). Boards and senior leaders should retain explicit accountability for strategic choices and ensure that human review has real authority, time, and competence. The objective is not to preserve human involvement symbolically but to create a collaboration system in which human and machine contributions remain distinguishable, testable, and mutually corrective.

CRITICAL DISCUSSION

From human versus AI to configuration fit

A major implication of the synthesis is that the conventional human-versus-AI comparison is theoretically weak for strategic management. Strategic decisions are rarely single acts of prediction. They are sequences of framing, search, interpretation, evaluation, negotiation, authorization, and implementation. Different stages create different opportunities for delegation. For example, AI may be superior at screening thousands of market signals but inferior at determining whether an apparently attractive opportunity fits a firm's identity, political constraints, or implementation capacity. Treating the whole decision as delegated or augmented therefore hides the actual architecture of collaboration. The more useful question is which actor should perform which component, in what sequence, with what information, and under what review rule. This position is consistent with organization design accounts of AI and with the automation-augmentation paradox, both of which imply that technology changes the division and coordination of work rather than merely replacing a decision maker (Puranam, 2021; Raisch & Krakowski, 2021; Shrestha et al., 2019). It also explains why apparently contradictory empirical findings can coexist: different studies may be evaluating different configurations while using the same broad label of human-AI collaboration.

Configuration fit also clarifies why standalone model accuracy is an incomplete criterion for strategic deployment. A highly accurate model can still produce poor team performance if users cannot recognize when it is wrong, if its explanations create misleading confidence, or if its recommendations crowd out independent judgment. Conversely, a somewhat less accurate model can be more useful when its errors are transparent and

complementary to human expertise. Bansal et al. (2021) demonstrate this distinction by showing that the best standalone AI need not be the best teammate, while Vaccaro et al. (2024) show that combined systems do not reliably outperform the better standalone actor. These findings challenge procurement and governance practices that prioritize benchmark accuracy without evaluating interaction quality. For strategic applications, evaluation should include calibration, error correlation, contestability, and the extent to which the system improves the final organizational decision. The unit of performance should therefore be the decision system, not the model in isolation.

Decision quality as a multidimensional outcome

A second implication is that strategic decision quality should be measured more broadly than predictive accuracy. Many AI studies rely on outcomes that are easy to quantify, such as classification accuracy, forecast error, task completion time, or agreement with expert ratings. These measures are useful but incomplete for strategy. A strategic decision may be analytically accurate yet poorly aligned with organizational capabilities, difficult to implement, ethically problematic, or vulnerable to environmental change. Decision quality therefore includes consistency and evidence use, but also strategic fit, robustness, novelty, accountability, and implementation feasibility. Doshi et al. (2025) demonstrate that generative AI can approximate expert rankings when evaluations are aggregated, yet this does not by itself establish that AI can determine which strategic commitment an organization should make. The distinction between evaluating alternatives and authorizing commitment remains important. Future empirical work should consequently use multidimensional outcome measures and follow decisions long enough to observe whether apparently high-quality choices produce resilient organizational outcomes.

This broader conception also changes how efficiency should be interpreted. AI may reduce the time required to analyze a decision, but faster convergence can be harmful if it suppresses exploration or causes teams to stop searching too early. Strategic value sometimes comes from maintaining productive disagreement, revisiting assumptions, or delaying commitment until uncertainty is better understood. Generative AI can increase the number of alternatives considered, yet it can also homogenize thinking if managers use similar models and prompts trained on overlapping information. The risk is not only factual error but convergence around plausible conventional answers. Human augmentation should therefore preserve cognitive diversity and encourage explicit counterfactual reasoning. Managers can ask AI to generate opposing cases, identify assumptions that would reverse a recommendation, and compare perspectives from different stakeholder positions. Such practices transform AI from a source of answers into an instrument for disciplined strategic inquiry.

The accountability problem in delegated strategy

Delegation creates a persistent accountability asymmetry. The more organizations rely on AI for analysis and recommendation, the easier it becomes for human actors to describe outcomes as consequences of the system rather than of managerial choice. Yet strategic accountability cannot be delegated in the same way as computation because organizations remain responsible for objectives, data selection, model procurement, deployment conditions, and final commitments. This asymmetry becomes particularly important when model outputs are opaque or when several actors contribute to the decision. Kellogg et al. (2020) show that algorithms can redistribute control within organizations, while Li et al. (2021) emphasize the role of senior governance actors in directing AI. The review therefore supports explicit accountability maps that identify who owns the decision, who validates the model, who monitors performance, and who has authority to suspend use. Without such clarity, delegation can create an accountability gap precisely where strategic consequences are greatest.

Accountability is closely related to contestability. A human-in-the-loop arrangement is meaningful only when the human has enough information, time, competence, and organizational authority to challenge the system. If approval is routine, heavily time pressured, or culturally discouraged, formal oversight can become ceremonial. This is particularly likely when AI outputs are presented with numerical precision or fluent explanations that imply greater certainty than the underlying evidence warrants. Lebovitz et al. (2022) show that professionals develop practices for engaging with opaque AI, illustrating that critical judgment is an active organizational accomplishment. Strategic governance should therefore preserve dissent channels and require documentation of material overrides or disagreements. The purpose is not to create bureaucracy around every AI output but to

ensure that consequential recommendations remain open to reasoned challenge. Contestability is thus a mechanism through which augmentation protects decision quality.

Organizational capability and the maturity of human-AI collaboration

The findings further suggest that firms differ in their ability to benefit from the same AI technology because collaboration depends on complementary organizational capabilities. Data quality, AI literacy, workflow design, governance, and learning routines determine whether models become reliable decision inputs or isolated experiments. Brynjolfsson et al. (2021) show that general-purpose technologies often require substantial complementary intangible investment before productivity gains appear. A similar logic applies to strategic AI. Organizations that purchase advanced systems without redesigning decision processes may add another information source without improving judgment. Saup et al. (2026) show that embedding generative AI into strategic decision systems requires movement from experimentation toward formal integration and institutionalization, supported by governance and socio-technical alignment. This evidence reinforces the review's conclusion that AI adoption and AI decision quality are different phenomena. Adoption can be rapid while decision-system maturity remains low.

Maturity also implies the capacity to learn from collaboration outcomes. Organizations should record when AI recommendations are accepted, modified, or rejected and compare these choices with later outcomes where feedback is available. Such records can reveal whether humans systematically override correct advice, whether particular models fail under identifiable conditions, or whether certain decision categories are poorly suited to delegation. Sturm et al. (2021) emphasize coordination between human and machine learning, which is essential because both sides of the collaboration must adapt. A mature system therefore treats every consequential decision as a potential learning event. This approach can gradually replace generic trust with evidence-based calibration. It can also prevent organizations from repeating the same failure mode after model updates, personnel changes, or environmental shifts. The result is a dynamic governance model in which collaboration rules evolve with accumulated evidence rather than remaining fixed after implementation.

Implications for future generative AI research

Generative AI makes these research priorities more urgent because model capabilities are broad, rapidly changing, and difficult to characterize with a single benchmark. Future studies should examine not only whether executives use generative AI but how interaction sequences shape strategic reasoning. Experimental designs could vary whether managers receive AI advice before or after forming an independent judgment, whether the model provides one recommendation or several competing scenarios, and whether uncertainty information is displayed. Field studies could compare strategic teams that use AI for option generation, evaluation, red teaming, or final recommendation. Longitudinal work is needed to determine whether repeated AI use improves managerial learning or produces deskilling and dependence. Recent evidence from strategic decision evaluation and executive integration provides a foundation for this agenda, but the field still needs common constructs and comparable outcome measures (Doshi et al., 2025; Srinivas & Chetan, 2026).

The most important future question may concern how organizations preserve human agency without wasting AI capability. Excessive caution can leave valuable prediction and search capacity unused, while excessive delegation can weaken judgment and accountability. Research should therefore test adaptive collaboration designs in which the level of AI autonomy changes with task structure, confidence, novelty, and consequence severity. Such systems could delegate routine analytical components automatically but trigger deeper human review when inputs fall outside familiar ranges or when recommendations involve irreversible commitments. This would operationalize the contingency framework developed in Figure 4 and move the literature beyond static categories. The broader theoretical opportunity is to understand strategic decision-making as a changing system of distributed cognition in which humans and AI jointly produce, evaluate, and contest representations of the future. That perspective is better suited to the emerging reality of AI-enabled strategy than either technological determinism or a purely human-centered account.

CONCLUSION

This systematic review examined how AI delegation and AI augmentation influence strategic decision quality. The evidence shows that neither collaboration mode is universally superior. Delegation can improve speed, consistency, search capacity, and reduction of some forms of human noise, but its benefits depend on task structure, reliable data, clear objectives, and effective monitoring. Augmentation can combine machine prediction with human context, causal reasoning, tacit knowledge, and accountability, making it more suitable for ambiguous and consequential strategic problems. Yet augmentation can also fail when humans anchor on AI, override useful outputs irrationally, or participate only as nominal approvers. The central conclusion is therefore contingent: strategic decision quality improves when the allocation of work between humans and AI matches the structure and uncertainty of the decision and is supported by calibrated trust and organizational governance.

The review also identifies a clear research gap around generative AI. Existing evidence demonstrates important productivity and evaluative capabilities, but strategic decision-making requires more than producing plausible text or ranking alternatives. Future research should examine how generative AI changes problem framing, option generation, executive attention, coalition formation, and accountability over time. Longitudinal field studies and field experiments are especially needed because laboratory tasks rarely reproduce the irreversibility and political complexity of strategic choices. Researchers should also compare delegation, augmentation, and staged hybrid architectures using common multidimensional measures of decision quality rather than productivity alone. A further priority is to investigate how model updates, organizational learning, and governance interact as AI capabilities evolve.

The review has several limitations. Database coverage and English-language selection may have omitted relevant studies, and the heterogeneity of research designs and decision-quality measures prevented statistical pooling. The rapid development of generative AI also means that evidence can become dated quickly. In addition, some insights are transferred from adjacent organizational and professional decision contexts because strategic-level empirical research remains comparatively limited. These limitations do not undermine the synthesis but indicate where caution is required. Overall, the review suggests that the future of strategic decision-making is unlikely to be characterized by either human control or machine autonomy alone. More promising is a deliberately designed division of cognitive labor in which AI expands analytical reach while humans preserve contextual judgment, contestability, and responsibility.

REFERENCES

1. Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.
2. Allen, R. T., & Choudhury, P. (2022). Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. *Organization Science*, 33(1), 149-169.
3. Baird, A., & Maruping, L. M. (2021). The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. *MIS Quarterly*, 45(1), 315-341.
4. Bansal, G., Nushi, B., Kamar, E., Horvitz, E., & Weld, D. S. (2021). Is the most accurate AI the best teammate? Optimizing AI for teamwork. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13), 11405-11414.
5. Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Artificial intelligence in organizations: Implications for information systems research. *Journal of the Association for Information Systems*, 22(2), 281-303.
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
7. Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
8. Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333-372.

9. Cao, G., Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2021). Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. *Technovation*, 106, 102312.
10. Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management Journal*, 41(8), 1381-1411.
11. Csaszar, F. A., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. *Strategy Science*, 9(4), 322-345.
12. Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayner, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality (Harvard Business School Working Paper No. 24-013).
13. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126.
14. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155-1170.
15. Doshi, A. R., Bell, J. J., Mirzayev, E., & Vanneste, B. S. (2025). Generative artificial intelligence and evaluating strategic decisions. *Strategic Management Journal*, 46(3), 583-610. <https://doi.org/10.1002/smj.3677>
16. Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data: Evolution, challenges and research agenda. *International Journal of Information Management*, 48, 63-71.
17. Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62-70.
18. Felin, T., & Holweg, M. (2024). Theory is all you need: AI, human cognition, and causal reasoning. *Strategy Science*, 9(4), 346-371.
19. Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Will humans-in-the-loop become borgs? Merits and pitfalls of working with AI. *MIS Quarterly*, 45(3), 1527-1556.
20. Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive challenges in human-artificial intelligence collaboration: Investigating the path toward productive delegation. *Information Systems Research*, 33(2), 678-696.
21. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577-586.
22. Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
23. Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown Spark.
24. Keding, C., & Meissner, P. (2021). Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions. *Technological Forecasting and Social Change*, 171, 120970.
25. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410.
26. Kim, H., Glaeser, E. L., Hillis, A., Kominers, S. D., & Luca, M. (2024). Decision authority and the returns to algorithms. *Strategic Management Journal*, 45(4), 619-648.
27. Krakowski, S., Luger, J., & Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, 44(6), 1425-1452.
28. Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what. *MIS Quarterly*, 45(3), 1501-1526.
29. Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126-148.
30. Lee, B. C., & Chung, J. (2024). An empirical investigation of the impact of ChatGPT on creativity. *Nature Human Behaviour*, 8(10), 1906-1914.
31. Li, J., Li, M., Wang, X., & Thatcher, J. B. (2021). Strategic directions for AI: The role of CIOs and boards of directors. *MIS Quarterly*, 45(3), 1603-1643.

32. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103.
33. Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, 46(3), 552-571.
34. Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192.
35. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
36. Puranam, P. (2021). Human-AI collaborative decision-making as an organization design problem. *Journal of Organization Design*, 10, 75-80.
37. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. *Academy of Management Review*, 46(1), 192-210.
38. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66-83.
39. Shrestha, Y. R., Krishna, V., & von Krogh, G. (2021). Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *Journal of Business Research*, 123, 588-603.
40. Simkute, A., Tankelevitch, L., Kewenig, V., Scott, A. E., Sellen, A., & Rintel, S. (2024). Ironies of generative AI: Understanding and mitigating productivity loss in human-AI interaction. *International Journal of Human-Computer Interaction*. Advance online publication.
41. Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118.
42. Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
43. Srinivas, R., & Chetan, S. S. (2026). Integrating artificial intelligence in strategic decision-making: Contexts for delegation and augmentation. *Group Decision and Negotiation*, 35, Article 59. <https://doi.org/10.1007/s10726-026-10016-x>
44. Sturm, T., Gerlach, J. P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., Nan, N., & Buxmann, P. (2021). Coordinating human and machine learning for effective organizational learning. *MIS Quarterly*, 45(3), 1581-1602.
45. Teodorescu, M. H. M., Morse, L., Awwad, Y., & Kane, G. C. (2021). Failures of fairness in automation require a deeper understanding of human-ML augmentation. *MIS Quarterly*, 45(3), 1483-1500.
46. Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14(3), 207-222. <https://doi.org/10.1111/1467-8551.00375>
47. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
48. Vaccaro, M., Almaatouq, A., & Malone, T. W. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8(12), 2293-2303.
49. von Krogh, G. (2018). Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing. *Academy of Management Discoveries*, 4(4), 404-409.
50. Wang, W., Gao, G., & Agarwal, R. (2023). Friend or foe? Teaming between artificial intelligence and workers with variation in experience. *Management Science*, 70(9), 5753-5775.
51. Sabz, A., Tabesh, P., & Yu, L. (2026). AI in strategic alliance formation: A framework for human-AI collaboration. *Journal of Business Strategy*, 47(2), 272-288. <https://doi.org/10.1108/JBS-01-2025-0027>
52. Saup, T.-O., Asghar, J., Kanbach, D. K., & Kraus, S. (2026). From pilots to decision systems: Embedding generative AI into strategic decision-making through a socio-technical and governance lens. *Journal of Decision Systems*, 35(1), 2597835. <https://doi.org/10.1080/12460125.2025.2597835>