STRATEGY SCIENCE

Modeling AI-Human Collaboration as a Multi-Agent Adaptation

Journal:	*Strategy Science*
Manuscript ID	Draft
mstype:	Special Issue: Can AI Do Strategy?
Keywords:	Simulations, Complex Adaptive Systems, Organizational Learning, Technology Strategy, Decision Making

Abstract:	We develop an agent-based simulation to formalize AI–human collaboration as a function of task structure, advancing a generalizable framework for strategic decision-making in organizations. Distinguishing between heuristic-based human adaptation and rule-based AI search, we model interactions across modular (parallel) and sequenced (interdependent) tasks using an NK model. Our results reveal that in modular tasks, AI often substitutes for humans—delivering higher payoffs unless human expertise is very high, and the AI search space is either narrowly focused or extremely broad. In sequenced tasks, interesting complementarities emerge. When an expert human initiates the search and AI subsequently refines it, aggregate performance is maximized. Conversely, when AI leads, excessive heuristic refinement by the human can reduce payoffs. We also show that even “hallucinatory” AI—lacking memory or structure—can improve outcomes when augmenting low-capability humans by helping escape local optima. These results yield a robust implication: the effectiveness of AI–human collaboration depends less on context or industry, and more on the underlying task structure. By elevating task decomposition as the central unit of analysis, our model provides a transferable lens for strategic decision-making involving humans and an agentic AI across diverse organizational settings.

"…Your old road is rapidly agin'
Please get out of the new one
If you can't lend your hand
For the times they are a-changin'…"
– Bob Dylan (The Times They Are A-Changin', 1964).

1. INTRODUCTION

Indeed, the times are changing—and fast. As AI and GenAI increasingly take on critical roles in organizations, the central question that confronts us is how humans and AI can best work together. Can AI and humans lend each other a hand, forging complementary collaboration, or is one destined to step aside to let the other take over depending upon the task? In this paper, we present a theoretical framework – using an agent-based simulation – that delineates the extent of AI-human complementarity based on task division within an organization, while identifying conditions where AI and humans may shift from collaboration to substitution.

Existing research on AI-human collaboration highlights AI's potential to automate and augment human capabilities in organizations (Csaszar, Ketkar, & Kim, 2024; Raisch & Krakowski, 2021). These complementarities have been explored through empirical studies, experiments (Choudhury et al., 2020; Zhang et al., 2022; Choi et al., 2025) or theorized based on the relative strengths of AI—such as predictive accuracy—versus contextual or domain expertise of humans (Choudhary et al., 2025; Shrestha et al., 2019). While invaluable, this work lacks a generalizable framework linking AI-human complementarity to organizational task division. A key reason for this gap is the prevailing view of AI as merely "an aid to prediction and decision-making" (Choudhary et al., 2025; p. 540), enhancing or automating human tasks (Csaszar, Ketkar, & Kim, 2024). However, scholars increasingly recognize AI as an agent in its own right—one that, like humans, can independently take strategic decisions and solve complex organizational problems (Krakowski, Luger, & Raisch, 2023; Acharya et al., 2025). This shift demands a re-evaluation of how AI fits into the broader fabric of organizational decision-making—no longer as an aid, but as a co-acting strategic agent. Therefore, we ask: *What makes AI and human agents distinct, and when do these differences create complementarities in organizational tasks?*

Adopting an agentic view of AI foregrounds a task-structure-based approach to analyzing AI-human collaboration—an essential level of aggregation in agentic decision-making that shapes the strategic allocation of resources and capabilities (Simon, 1955; Zenger, 2002; Nickerson, Silverman, & Zenger, 2007; Levinthal &

Wu, 2024). Accordingly, we model AI-human interaction as a multi-agent system engaged in joint problem-solving, highlighting the conditions under which complementarities emerge—and where AI and humans may become substitutes depending on the structure of task division, such as modular versus sequenced tasks (Raisch & Fomina, 2023). The foundation of our model lies in the distinct attributes of AI and Human (henceforth, H) – the two agents in our conceptualization – in adaptive search, characterized by two classical parameters of the NK model (Levinthal, 1997; Adner, Csaszar, & Zemsky, 2014): the search space (N parameters) and the complexity of the search landscape (K), which determine how adaptive search unfolds for an agent. In our framework, H is assumed, as per convention, to be boundedly rational (Puranam, Stieglitz, Osman, & Pillutla, 2015; Simon, 2000). Therefore, H operates within a *narrow search space* and relies on a *heuristic-based approach* (Gigerenzer, 2004; 2020; Levinthal & March, 1981)—such as giving higher weights to most recent decisions states—when probabilistically determining the next state in N, for a given K (Leiblein & Macher, 2009).

In contrast, we relax the bounded rationality assumption for AI (Nishant et al., 2024; Shick et al., 2024), allowing it to explore a *larger search space* (Raisch & Fomina, 2023) and apply a *rule-based approach* to updating its search, even for high values of K.[1] The overall payoff function in our model is computed as the average of the payoffs from AI's and H's search processes in collaboration. For sequential tasks, whether AI-to-H or H-to-AI, we introduce an additional parameter, C, upgrading the NK model to its NKC representation (Gavetti et al., 2017; Biggiero, 2016; Suzuki & Arita, 2005). Here, C captures coevolution of the second agent in the sequence, allowing it to refine its search based on the states explored by the first agent. This adjustment influences the second agent's search process within the landscape. It follows that, for modular tasks, C=0.

Our simulations reveal that for modular tasks (C=0), AI's search and adaptation process often diminishes the marginal value of human (H) domain expertise, suggesting that investing in advanced AI may yield higher returns, on average, than focusing on human training. However, our findings also indicate specific regions of AI-H complementarity even within modular tasks. When *H lacks the expertise* to evaluate complexities (low K), the optimal outcomes occur when the AI's search space is moderately vast. If the search space is too

[1] This formalization does not imply full rationality for AI; instead, it differentiates between bounded rationality and computational boundedness (Wheeler, 2018). Unlike cognitively bounded humans, AI's limitations are primarily computational (Nishant, Schneckenberg, & Ravishankar, 2024).

narrow, both AI and H risk getting trapped in sub-optimal local peaks, diminishing the overall payoff. Conversely, if the AI's search space is too expansive, its computational advantages may be underutilized as H fails to contribute meaningfully, thereby reducing the average payoff. Crucially, our model highlights two scenarios where *H's high domain expertise* adds substantial value. First, focused tasks where the AI does not need to explore a vast space, allowing H's deep expertise to enhance performance. Second, a division of labour between AI and H, where AI efficiently scans a broad landscape—such as screening M&A deals using extensive historical financial data—while H independently applies experiential knowledge to a more focused M&A evaluation, for instance, assessing CEO's psychological attributes as opposed to more standardized financial analysis that AI can perform.

For sequenced tasks ($C > 0$), we analyze two task flow sequences to derive insights. In AI-to-H sequencing, we find that a moderately high level of domain expertise in H (i.e., a moderately high C) serves as the most resilient complement to AI's adaptive search in Step 1. By observing AI's strategic landscape, H can refine their own search process through heuristic-based updates, effectively channelling AI's exploration—even in vast search spaces. However, as H incorporates an excessive number of parameters from AI's search (i.e., when C is very high), the absolute payoff and its resilience across the search space decline rapidly. Applying a heuristic-based search over too many parameters increases the risk of cognitive biases that offset the benefits of AI's rule-based search in Step 1, which can rapidly compound sub-optimal payoffs associated with poor strategic choices. The managerial takeaway for harnessing complementarity in AI-to-H sequenced tasks is clear: excessive human intervention is suboptimal. Instead, the optimal strategy is to balance AI's breadth of rule-based search with H's domain expertise through moderate calibration using suitable heuristics.

Second, in the H-to-AI sequencing, we find that optimal performance arises when a rule-based AI calibrates the prior search of a highly capable H agent. Among all possible sequencing configurations, this setup maximizes joint payoff. Building on this key insight, we statistically generalize the following result: all else being equal, why is an H-to-AI sequence—where H is highly capable—preferable to an AI-to-H sequence as a task structure? We formalize this argument in the discussion section.

Striking insights emerge when we distinguish between rule-based and "hallucinatory AI" (which operates without memory and updates states randomly) in Step 2. In a moderately vast AI search space, even a hallucinatory AI performs nearly as well as a rule-based AI in augmenting H's prior search. However, while the benefits of rule-based AI remain stable, the effectiveness of hallucinatory AI declines sharply as the search space expands. Perhaps the most counterintuitive yet striking insight is this: hallucinatory AI outperforms rule-based AI when augmenting the prior search of an H agent who is not highly capable or lacks domain expertise. Moreover, in this paradigm (low-capability H), the benefits of hallucinatory AI are more resilient than those of rule-based AI. This resilience likely stems from AI's memory-less hallucination, which helps pull H out of local traps formed in Step 1. In contrast, a rule-based AI retains some memory of H's previous search. If the low-capability H was stuck on a suboptimal peak, this memory can reinforce localized search by a rule-based AI, deepening entrenchment in a suboptimal outcome. However, broader implications of this insight warrant discussions on AI ethics, trust, and regulation. After all, how much can one trust a hallucinating agent if it lacks robustness to ethical and reliability standards?

We make two contributions. First, by foregrounding task structure as the central unit of analysis, this paper offers a generalizable framework to study AI–human collaboration in strategic decision-making. We formalize the contrast between agentic AI—capable of rule-based, broad search—and boundedly rational humans—engaging in heuristic-based, narrow search—within an NKC simulation model (Knudsen, 2024; Gavetti, Helfat, & Marengo, 2017; Biggiero, 2016). This allows us to analyze how collaboration unfolds across modular and sequenced tasks (Raisch & Fomina, 2023), and under what conditions AI complements or substitutes for human decision-making.

Second, our results speak to foundational ideas in strategy around modularity, decomposability, and technology architecture, and extend them in the context of AI-human interactions (Agarwal & Kapoor, 2023; Burton-Jones et al., 2021; Baird & Maruping, 2021). In modular tasks, AI frequently substitutes for humans, consistent with prior work showing decomposable systems favor automation (Kapoor & Adner, 2012; Kapoor, 2013; Yayavaram & Ahuja, 2008; Birhane, 2021). In sequenced tasks, we find the strongest complementarities emerge when expert humans lead, and AI refines—a counterintuitive finding that challenges the dominant view that AI should act first as a filter with humans adding nuance later. Another surprising result is that even a hallucinatory AI—one that updates its decision space in a memory-less manner—can outperform rule-based AI when paired with low-capability humans. Together, these insights advance the literature on AI–human complementarity in strategic decision-making by giving managers a framework to evaluate how agentic configurations and task structures jointly shape performance in organizational tasks that will inevitably feature AI-human interactions across industry contexts.

2. BACKGROUND LITERATURE

2.1. AI-Human Collaboration

Given the rapid rise of AI and GenAI in organizational processes, a growing research stream in management science has explored AI-human collaboration, highlighting AI's potential to both automate and augment human capabilities (Csaszar, Ketkar, & Kim, 2024; Raisch & Krakowski, 2021). These complementarities have been examined through empirical studies and experiments within specific organizational contexts (Choudhury et al., 2020; Zhang et al., 2022; Choi et al., 2025) and theorized based on AI's relative strengths—such as its vast data processing capacity and predictive accuracy—compared to humans' contextual and domain expertise (Choudhary et al., 2025; Shrestha et al., 2019).

Much of the literature has emphasized AI as a tool for enhancing or automating human tasks, viewing it primarily as an aid to prediction and decision-making (Choudhary et al., 2025, p. 540; Csaszar, Ketkar, & Kim, 2024). However, a growing body of work now recognizes AI as an agent in its own right, capable of independently solving complex organizational problems (Krakowski, Luger, & Raisch, 2023; Acharya et al., 2025). This shift—from AI as a tool to AI as an agent—is a critical one, as it moves beyond mere augmentation of human processes to potentially reshaping human decision-making itself (Krakowski, Luger, & Raisch, 2023; Hillebrand, Raisch & Schad, 2025). In this agentic paradigm, a fundamental question emerges—one that this paper seeks to address: What makes AI and human agents distinct, and when do these differences create complementarities in organizational tasks?

Related literature in Information Systems (IS) has viewed AI-human complementarity from a *systems perspective* using an agentic conceptualization of AI. Burton-Jones et al. (2021) challenge the assumption of human primacy in IS use, arguing that AI-driven systems increasingly act as autonomous agents rather than passive tools. They introduce delegation as a key lens, where humans and AI transfer decision-making authority dynamically to solve system-level tasks. Wang (2021) highlights how AI-enabled system architectures combined with humans' strategic oversight improves systems performance. Baird and Maruping (2021) explore AI delegation, showing AI as an agentic entity capable of autonomous process innovations. Together, these perspectives position AI as agentic but largely to drive efficiency in processes, and not decision-making.

In the field of strategic management, Sebastian Raisch and their peers have significantly contributed to shifting the focus to decision-making by adopting a *task-structure* level de-construction of the phenomenon. First, they distinguish between automation and augmentation (Raisch & Krakowski, 2021), demonstrating that while AI as a tool enhances efficiency through automation, AI as an agent necessitates structured human-AI collaboration to fully leverage augmentation benefits. Second, they examine the efficiency gains within the automation paradigm, arguing that AI-driven decision-making can both substitute for and complement human managerial tasks, thereby reshaping competitive advantage (Krakowski, Luger, & Raisch, 2023). Third, Hillebrand, Raisch, and Schad (2025) introduce a multilevel perspective, shifting the focus from isolated task-level applications to broader organizational structures, emphasizing the role of collective agency of AI and humans. Finally, within this agentic view of augmentation, Raisch and Fomina (2023) identify distinct AI-human hybrid problem-solving processes within modular and sequential organizational task structures. Their key insight is that modular search by AI facilitates distant exploration, while AI-augmented sequential search enhances the quality of local solutions. Building on this conceptual foundation, our work extends and models the contingencies of AI-human hybrid problem-solving across diverse organizational task structures, moving beyond the sole focus on the relative breadth of human and AI search spaces (Raisch & Fomina, 2023).

2.2. Agentic AI and Organizational Task Structure

The structure of organizational tasks represents the most fundamental level of aggregation of agentic decisions, and it plays a critical role in shaping the strategic allocation of resources, capabilities, and attention within organizations (Simon, 1955; Zenger, 2002; Puranam, 2018; Levinthal & Wu, 2024). Despite its foundational importance in theories of organizational design and decision-making, the discourse on AI-human collaboration has, to date, largely overlooked the task structure as a locus of analysis. This oversight may stem from the implicit tendency to treat AI primarily as a tool rather than as an agentic actor. While recent work (e.g., Raisch & Fomina, 2023) has reintroduced the relevance of task design in the context of AI-as-tool, the agentic potential of AI remains underexplored (Vanneste & Puranam, 2024; Kim, Khoreva, & Vaiman, 2024; Feuerriegel et al., 2024). Therefore, before presenting our formal model, it is essential to highlight the salience of a task-structure-based approach as a generalizable organizational lens for understanding AI-human collaboration.

Various perspectives have attempted to categorize the division of labor between AI and human agents, much like how organizational theorists traditionally conceptualized division of labor between any two human agents—exploration versus exploitation (Wagner, 2020; Lazer & Friedman, 2007; Luger, Raisch, & Schimmer, 2018) specialization versus generalization (Weiss, 1999; Choudhary et al., 2025; Choudhury et al., 2020), low versus high uncertainty (Felin & Holweg, 2024), and so on. While each of these perspectives offers valuable insights, they are limited in their ability to account for AI's role not just as a tool, but as an autonomous decision-making agent (Shrestha, Ben-Menahem, & Von Krogh, 2019). This approach certainly clarifies organizational processes ripe for AI adoption but does not account for the iterative interactions in joint decision-making by AI and humans. The missing link is a perspective that accounts for the fundamental task structure in which AI and humans engage in joint problem solving and decision making.

In sum, analyzing AI's role from a top-down perspective—beginning with task structures, followed by context-specific systems and their components—reveals AI primarily as a tool enhancing efficiency within predefined human workflows. Conversely, adopting a bottom-up perspective—starting with components, moving upward through systems, and ultimately addressing the broader task structure—reveals AI taking on a more agentic role, actively shaping fundamental organizational decisions (Bader & Kaiser, 2019). This shift toward a task-structure-based perspective, which strategically informs organizational resource and capability allocation decisions (Levinthal & Wu, 2024; Aggarwal & Wu, 2015), is particularly critical. While recent scholarship acknowledges the importance of analyzing AI's agentic potential from a task-structure standpoint (Raisch & Fomina, 2023; Shrestha et al., 2019), formal modeling of these dynamics remains unexplored. Figure

1 schematically illustrates the significance of employing a task-structure-based approach to understand and evaluate AI-human collaboration in organizations.

---Insert Fig. 1 here---

3. MODEL IMPLEMENTATION

Building on the premise that the underlying task structure of an organizational system is the most suitable level for analyzing AI-human collaboration (Raisch & Fomina, 2023), we leverage the widely recognized NK model, which has been used to study adaptive search in modular and sequenced task environments (Fang & Kim, 2018). Our approach introduces two key refinements to the traditional NK models of adaptive human search (Levinthal, 1997; Rivkin & Siggelkow, 2003; Knudsen & Srikanth, 2014). First, we vary both N (the search space) and K (the complexity of search) between human and AI agents to reflect differences in bounded rationality and search methodologies—rule-based for AI versus heuristic-driven for humans. This distinction allows us to capture the fundamental ways in which these agents – AI and human – approach problem-solving.

Second, we introduce a parameter, C, which is particularly relevant for sequential tasks. This parameter enables AI and H to coevolve, adapting their search strategies based on prior searches conducted by the other. This extension aligns with the NKC model, a variant of NK models used to study coevolutionary dynamics that shape adaptive search (Gavetti et al., 2017; Biggiero, 2016; Suzuki & Arita, 2005).

3.1. Model Parameters

3.1.1. Search space, N

As per convention, the parameter N represents the number of decision variables available to an agent, capturing the expanse of its search space (Kauffman, 1993, 1995). Each decision variable within this search space is updated probabilistically, assuming a binary state of either 0 or 1 (drawn from a Bernoulli distribution), following traditional NK models used in organizational studies (Levinthal, 1997; Rivkin & Siggelkow, 2003; Adner, Csaszar, & Zemsky, 2014). Therefore, the AI and human (H) search spaces are represented by the following sets of decision variables:

$$N_{AI} = \{x_1^{AI}, x_2^{AI},..., x_N^{AI}\}, \text{ and}$$

$$N_H = \{x_1^{H}, x_2^{H},..., x_N^{H}\};$$

where $x_i^{AI/H}$ can take the decision state of either 0 or 1.

The AI agent, owing to its computational capacity for high-dimensional search, is assigned a broader search space than the boundedly rational H agent. To formalize this assumption, we impose the constraint on the cardinalities of the two sets, $|N_{AI}| > |N_H|$, ensuring that AI agent always evaluates a larger set of decision variables than its H counterpart in our model (Raisch & Fomina, 2023).

3.1.2. Search complexity, K

The complexity of the search process is governed by the parameter K, which reflects the degree of interdependence among decision variables. A higher K results in a more rugged search landscape, meaning a greater number of local peaks in the payoff function, making adaptive search for a global optimum more challenging (Levinthal, 1997). When K = 0, all decision states in N are independent, leading to a smooth landscape with a single global optimum. In our model, K_{AI} represents the AI's ability to recognize interdependencies among decision variables in N_{AI}, while K_H captures the H agent's capacity to do the same within its own search space, N_H. As per standard practice, for a given N, K can take values from 0 (smooth landscape) to N-1 (extremely rugged landscape) (Adner, Csaszar, & Zemsky, 2014).

3.1.3. Co-evolution strength, C

The parameter C represents the strength of interdependence between AI and H's search process. In the case of modular tasks (C=0), both evolve independently. In the case of sequential tasks (C>0) one agent's decisions shapes the other's adaptive search. This leads us to model two configurations: AI-to-H, where AI's search adapts to H search, and H-to-AI, where H decisions guide AI adaptation. In both cases, the magnitude of C determines how tightly the two agents, AI and H, coevolve to derive a joint payoff from the adaptive search.

3.2. Modelling Adaptation

By defining K, we establish a fundamental distinction between heuristic-based human search and rule-based AI search in determining the probabilistic state of the next decision variable within N. This process builds upon a previously identified subsequence of decision states, $S_i \subseteq N$ represented as $\{x_1, x_2, ..., x_i\}$, which reflects the adaptive search up to that point. While the first key distinction between the AI and H search spaces is that

$|N_{AI}| > |N_H|$, the second critical difference lies in how we model their search approaches: heuristic-based for H and rule-based for AI, to determine the probabilistic decision state of x_{i+1}, for a given K.

3.2.1. Heuristic-Based Adaptation for H

Human decision-making is inherently heuristic-based, relying on cognitive shortcuts, contextual knowledge, and common sense, rather than exhaustive evaluation of all available options (Gigerenzer, 2004; 2020; Levinthal & March, 1981; Nishant, Schneckenberg, & Ravishankar, 2024). We model this heuristic-based adaptation process using a weighted function, where past decision states influence the decision on the current decision parameter with varying degrees of emphasis (Leiblein & Macher, 2009). The heuristic-based adaptation by H to determine a subsequent decision state is defined as:

$$x_{i+1}^{H} = \begin{cases} 1, & if \ \dfrac{\sum_{j=1}^{K_H} j * x_{i-j}^{H}}{\sum_{j=1}^{K_H} j} \geq 0.5, \\ 0, & otherwise. \end{cases}$$

where *j* represents the weight assigned to past decisions, ensuring that more recent decisions exert a stronger influence on the adaptation process. This formulation captures the tendency of human agents to reinforce recent patterns while averaging historical data, aligning with established models of strategic adaptation under uncertainty (Rivkin & Siggelkow, 2003; Levinthal, 1997).[2]

The relative benefits and costs of a heuristic-based adaptation is not obvious. For instance, while determining interdependencies between only a few decision states (i.e., K_H is *low)*, heuristic-based adaptation can reflect H's contextual updating of priors. In contrast, when heuristics are applied to a broader set of decision parameters (i.e., K_H is *high)*, the same adaptive method can manifest in the form of bias where excessive emphasis is placed on recent decision states while appropriating relatively lower emphasis to historical decisions. Therefore, heuristic-based adaptation by H can either enhance efficiency in the search process by contextual fine-tuning of recent decision states, or hinder exploration by trapping the search process in local optima due to excessive recency bias.

[2] For simplicity, we assume a weighting scheme, where recent decision states receive progressively (linearly) higher weights compared to earlier ones. This assumption reflects the bounded rationality inherent in human decision-making, where cognitive factors increase tendency in humans to give a higher relative weightage to recent decisions.

3.2.2. Rule-Based Adaptation for AI

Unlike humans, we model AI to conduct rule-based adaptive search that evaluates all prior decision states with equal weighting to determine the subsequent state (Nishant, Schneckenberg, & Ravishankar, 2024; Shick et al., 2024).[3] This distinction underscores a fundamental principle of AI decision-making—rather than heuristic-based recalibrations, AI follows a predefined rule to adapt its search process, that is modelled as:

$$x_{i+1}^{AI} = \begin{cases} 1, & if\ \frac{1}{K_{AI}} \sum_{j=1}^{K_{AI}} x_{i-j}^{AI} \geq 0.5, \\ 0, & otherwise. \end{cases}$$

where x_{i+1}^{AI} represents the binary state of the AI decision variable. Since all prior decisions states are equally weighted (rule-based), AI decision state updating follows a simple averaging mechanism. Figure 2 provides a stylized example to distinguish between heuristic and rule-based adaptive search for conceptual clarity.

--Insert Fig. 2 here--

3.3. Modular Tasks

In modular task search, the configuration resembles traditional agent-based simulations (Kauffman, 1993, 1995; Levinthal, 1997), but in a setting where AI and H operate in parallel, independently performing adaptive search following rule and heuristic-based adaptation, respectively. We define the pay-off function as the average of the individual pay-off functions of AI and H searches, each operationalized as per conventional practice in NK models. The individual payoffs for AI and H are defined as follows:

For the AI agent:

$$PO_{AI} = \frac{1}{|N_{AI}|} \sum_{i=1}^{N_{AI}} \left(x_i^{AI}\right)$$

For the H agent:

[3] By rule-based, we refer to AI's "mechanical" mode of decision-making, grounded in the "formal rationality of mathematical optimization procedures (Nishant, Schneckenberg, & Ravishankar, 2024, p. 19)." It is important not to conflate rule-based adaptation logic with unbiasedness. Rule-based adaptation can characterize both biased and unbiased AI systems, as biases embedded in training data can carry over into out-of-sample predictions—even when the underlying decision process remains formally rule-driven.

$$PO_H = \frac{1}{|N_H|} \sum_{i=1}^{N_H} (x_i^H)$$

It's important to note that these pay-offs correspond to a specific realized state of N_{AI} and N_H. However, in NK models, the fitness contributions ($x_i^{AI/H}$) are probabilistically determined. Following established convention (e.g., Clement & Puranam, 2018), we therefore compute the pay-offs from AI and H searches by averaging PO_{AI} and PO_H over 1,000 simulated runs of the model. So, what we analyze is actually $\overline{PO_{AI}}$ and $\overline{PO_H}$ over 1000 runs, for a given combination of N and K.

The joint payoff from AI and H's independent adaptive search, APO, is simply calculated as:

$$APO = \left[\frac{\overline{PO_{AI}} + \overline{PO_H}}{2}\right]$$

Following standard practice, the implicit assumption is that the decision state '0' maps to a 'bad pay-off', and the decision state '1' maps to a 'good pay-off' (Puranam et al., 2015). So, 0 < $\overline{PO_{AI/H}}$ < 1, and 0 < APO < 1.

3.4. Sequenced Tasks

3.4.1. AI-to-Human (ATH)

In this task structure, AI initiates search, generating a structured sequence of decision states that subsequently guides H's search. AI updates its decision sates through a rule-based adaptive approach as discussed earlier. Once AI has generated its sequence N_{AI}, a subset of C (where C < $|N_{AI}|$) decision states, starting from the first, serves as the initial sequence of decision states for H to begin adaptive search. This subset is denoted as:

$$S_H = \{x_1^{AI}, x_2^{AI}, \ldots, x_C^{AI}\}$$

Unlike in the case of modular tasks, H *does not* initiate adaptation from a random distribution; rather it adapts to a predefined sequence, making H's adaptation path-dependent on prior AI-generated decision states. The parameter C effectively captures the level of dependency of H's adaptive search on the decision states of N_{AI} to generate its own decision states in N_H. For a given value of C, decision states in N_H are generated through the heuristic-based updating applied on S_H.

3.4.2. Human-to-AI (HTA)

In this task configuration, H initiates the search, generating a structured sequence of decision states following heuristic-based adaptation which subsequently guides AI's search. H's heuristic-based search produces a structured sequence of decision states, N_H, from which a subset of C values serves as the initial seed for AI's adaptive search:

$$S_{AI} = \{x_1^H, x_2^H, \dots, x_C^H\}, \qquad where\ 1 \leq C < |N_H|$$

Therefore, S_{AI} represents a pre-structured H subspace that governs how AI decision states evolve. Unlike modular tasks, where AI starts search using a random draw, here AI starts adapting its own search by using rule-based adjustments to S_{AI}, making its adaptive search dependent on prior search by H.

3.4.2.1. Memory Fading of H Search

The transition from H-guided initialization (S_{AI}) to AI-driven perpetuation follows a progressive "memory-fading" mechanism. AI starts by applying rule-based search on S_{AI} (for a given C) but, unlike the case of AI-to-H sequence, AI runs out of H's decision states to apply its rule-based adaptation to beyond a point because $N_{AI}/N_H > 1$ by design. Therefore, after N_H is fully exhausted, AI starts to perpetuate the remainder of its search to complete strategic explorations on N_{AI} using information from its own realized decision states. Therefore, as the AI search space expands, memory of H's prior search starts fading gradually. To illustrate this complex transition, consider the stylized example where $N_H = 6$, $C = 4$, and $N_{AI} = 8$. Assume the following H sequence of decision states (after Step 1 search by H) which acts as the initial seed for AI updating:

$$S_{AI} = [1, 0, 1, 1, 0, 1]$$

AI then updates its sequence using the memory-fading rule-based approach as follows:

Element in N_{AI}	AI sequence dependence	Source	Rule-based calculation	AI decision state (x^{AI})
x_1^{AI}	$f([1, 0, 1, 1])$	S_{AI}	(1+0+1+1)/4 = 0.75	1
x_2^{AI}	$f([0, 1, 1, 0])$	S_{AI}	(0+1+1+0)/4 = 0.5	0
x_3^{AI}	$f([1, 1, 0, 1])$	S_{AI}	(1+1+0+1)/4 = 0.75	1
x_4^{AI}	$f([1, 0, 1, x_1^{AI} = 1])$	S_{AI} + AI-generating sequence	(1+0+1+1)/4 = 0.75	1
x_5^{AI}	$f([0, 1, x_1^{AI} = 1, x_2^{AI} = 0])$	S_{AI} + AI-generating sequence	(0+1+1+0)/4 = 0.5	0
x_6^{AI}	$f([1, x_1^{AI} = 1, x_2^{AI} = 0, x_3^{AI} = 1])$	S_{AI} + AI-generating sequence	(1+1+0+1)/4 = 0.75	1

At this stage (i.e., by the time the 6th decision state has been realized in N_{AI}), the memory of H's search has completely faded. Once the AI sequence reaches $|S_{AI}| = |N_H|$, we model AI's adaptation using two approaches: a) Rule-based adaptive search, and b) hallucinatory exploration (random; no rules or heuristics). We describe the nuances of these distinct approaches for AI's search perpetuation below.

3.4.2.2. Rule-based AI perpetuation

In rule-based perpetuation, AI continues generating decision states in N_{AI} following the same process as in AI's modular search or Step 1 of AI-to-H sequenced task. Effectively, C starts to function as K_{AI}, to calculate the subsequent decision state as follows:

$$x_{i+1}^{AI} = \begin{cases} 1, & if\ \frac{1}{C}\sum_{j=1}^{C} x_{i-j}^{AI} \geq 0.5, \\ 0, & otherwise. \end{cases}$$

Continuing with our stylized example, for N_{AI} = 8, the AI sequence extends as follows:

Element in N_{AI}	AI sequence dependence	Source	Rule-based calculation	AI decision state (x^{AI})
x_7^{AI}	$f([x_3^{AI} = 1, x_4^{AI} = 1, x_5^{AI} = 0, x_6^{AI} = 1])$	AI-generating sequence	(1+1+0+1)/4 = 0.75	1
x_8^{AI}	$f([x_4^{AI} = 1, x_5^{AI} = 0, x_6^{AI} = 1, x_7^{AI} = 1])$	AI-generating sequence	(1+0+1+1)/4 = 0.75	1

3.4.2.3. Hallucinatory AI perpetuation

Alternatively, we model AI such that it completely abandons structured adaptation and transitions into hallucinatory search, where all remaining decision states are generated randomly:

$$x_{i+1}^{AI} \sim Bernoulli\ (0.5), \forall\ i \geq |N_H|.$$

The remaining decision states in N_{AI} = 8 would now be generated as follows:

Element in N_{AI}	AI sequence dependence	Source	Rule-based calculation	AI decision state (x^{AI})
x_7^{AI}	Randomly assigned (0 or 1)	~ Bernoulli (0.5)	None	0 / 1 (Random)
x_8^{AI}	Randomly assigned (0 or 1)	~ Bernoulli (0.5)	None	0 / 1 (Random)

3.4.3. Payoff Evaluation in Sequenced Tasks

The payoff evaluation in sequenced tasks follows the same formulation as the modular case, with the aggregated system payoff, APO, derived from individual payoffs, $\overline{PO_{AI}}$ and $\overline{PO_H}$. Unlike modular tasks where $C=0$, APO in sequenced tasks is a function of $N_{AI/H}$, $K_{AI/H}$ and C (where $C > 0$).

4. MODEL RESULTS

We now present the simulation results across the three modeled configurations: a) modular tasks with independent search (C=0), b) AI-to-H sequential search, and c) H-to-AI sequential search. Each section discusses the conditions under which agent complementarities maximize joint payoff and highlights key task-structure contingencies.

4.1. Modular tasks

We simulate and plot the average pay-off from independent modular searches conducted by a human agent (H) and an AI agent, under varying conditions of their relative ability to recognize interdependencies among decision variables—captured by the ratio K_H/K_{AI}. This ratio reflects H's decision-making sophistication relative to AI attributable to H's domain expertise. Our first key finding is that, regardless of the K_H/K_{AI} ratio, the joint pay-off follows a curvilinear relationship with the relative size of the agents' search spaces—represented by N_{AI}/N_H.[4] When AI's search space is either too small (comparable to H's) or excessively large, the joint pay-off from modular search declines. The optimal outcome emerges when AI operates in a moderately expansive search space, with $N_{AI}/N_H = 5$.

The highest pay-off occurs when this moderate AI search space ($5 < N_{AI}/N_H < 6$) is paired with low sophistication in H's adaptive search (i.e., $K_H/K_{AI} < 0.5$). This insight is central to understanding the value of modular task configurations: when tasks are modularized and executed independently, organizations benefit most when the AI is relatively advanced, and H avoids overcomplicating their search through excessive heuristic-based adaptation. In essence, modular search is most effective when AI applies rule-based search for moderately higher interdependencies among its decision variables—about twice that of H's.

[4] Technically, the ratio of the cardinalities of the two sets, N_{AI} & N_H, as described in the model setup earlier.

Does this imply that H is redundant in modular settings? Not necessarily. When H possesses high expertise—capable of identifying many interdependencies (i.e., $K_H/K_{AI} > 3$)—there are still configurations that yield strong outcomes, though the average pay-off is lower than in the low K_H/K_{AI} case. Specifically, expert H search enhances performance under two conditions: when the AI search space is too narrow (low N_{AI}/N_H) or when the AI's search space is overwhelmingly large (extremely high N_{AI}/N_H). While the former is intuitive, the latter illustrates a powerful form of AI-H complementarity. Here, the heuristic-based adaptation of an expert H is effectively complemented by the rule-based search of an advanced AI operating over a vast decision space. Figure 3 illustrates the joint payoffs from AI and H under varying modular search conditions.

--Insert Fig. 3 here--

4.2. Sequenced task: AI-to-H

In the first sequence, AI initiates the process by applying rule-based adaptation across its search space (N_{AI}). This is followed by H's heuristic-based adaptation, which refines AI's output. The extent of this refinement is governed by the co-evolution parameter $C > 0$. Specifically, C indicates the number of AI-derived decision states on which H performs its own heuristic-based adaptive search. Our findings show that the joint pay-off generally follows a curvilinear relationship with the N_{AI}/N_H ratio. This suggests that when AI operates within a moderately expansive search space, the combination of AI's rule-based approach and H's heuristic-based refinement complement one another to yield maximum benefits. However, these benefits diminish when H applies heuristics too aggressively after AI's rule-based search—particularly when the ratio C/K_{AI} exceeds 0.5 or, more strongly, exceeds 2. This effect is especially pronounced when AI's initial search space is extremely vast ($N_{AI}/N_H > 5$). In such cases, excessive heuristic refinement by H ($C/K_{AI} > 2$) actually reduces overall pay-off rapidly. Thus, in AI-to-H sequences, the best outcomes are achieved when a broad initial search by AI is followed by a moderate degree of heuristic-based refinement of the decisions by H. Figure 4 illustrates the joint payoffs from AI-to-H sequenced tasks.

--Insert Fig. 4 here--

4.3. Sequenced task: H-to-AI

In the second sequence, the process begins with H applying heuristic-based adaptation across its own search space (N_H). AI then follows by refining H's output through rule-based adaptation. In this setup, the parameter C represents the number of decision states produced by H on which AI performs its rule-based search. Unlike the AI-to-H sequence—where H had access to AI's full (and larger) search space during its adaptation (since $N_{AI}/N_H > 1$ by design)—in the H-to-AI sequence, AI is constrained to refining H's limited search output. This introduces two distinct sub-scenarios: a) *Rule-based perpetuation:* When C exceeds the size of H's search space ($C > |N_H|$), AI continues to search using its usual rule-based approach to adapt where K_{AI} effectively becomes equal to C, b) *Hallucinatory search:* Alternatively, when $C > |N_H|$, AI may perform a search decoupled from H's outputs—akin to hallucination—where it conducts random exploration across its vast space to update its decision states without relying on prior structure. We plot the aggregate pay-off as a function of the N_{AI}/N_H ratio under these two conditions: rule-based perpetuation and hallucinatory (random) search.

Our simulations reveal a few key insights. First, AI significantly enhances overall performance when it follows a high-capability H—that is, when H's Step 1 search yields a high pay-off. In these cases, rule-based AI consistently outperforms hallucinatory AI. Moreover, as AI's search space becomes exceptionally large ($N_{AI}/N_H > 15$), the diminishing returns from hallucinatory AI are sharp and rapid. However, a counterintuitive result emerges in cases where H has low capability (i.e., H's pay-off in Step 1 of search is low). In such scenarios, hallucinatory AI marginally outperforms rule-based AI, and this advantage persists even as AI's search space grows. The intuition here is striking: a hallucinatory AI, akin to random coin tossing, can help a low-capability H escape local optima. In contrast, a rule-based AI is constrained by the suboptimal structure inherited from H's poor search, thereby perpetuating its limitations.

Thus, in H-to-AI task sequences, two effective AI-H collaboration strategies emerge: a) Use rule-based AI to build upon and extend the high-quality search of an expert H, or b) Use hallucinatory AI to rescue or diversify from the limited search of a low-capability H, especially when AI has access to a large search space of decision parameters. Figure 5 illustrates the joint payoffs from H-to-AI sequenced tasks.

---Insert Fig. 5 here---

5. DISCUSSION

5.1. Managerial Implications

This paper formalizes how agentic AI systematically differs from human agents and explores the implications for AI-human collaboration in joint problem-solving. Using a simple yet fundamental NKC framework, we model AI's broad rule-based search alongside human heuristic-based search in modular and sequenced task structures. This provides a generalizable framework for AI-human complementarity, moving beyond industry- or systems-specific AI adoption, automation, and augmentation debates.

Our simulation results have important managerial implications. First, the relative emphasis on AI versus human both from a strategic and investment perspective should be guided by task structure rather than industry. As a thought experiment, imagine giving a researcher only the task structure of a Fortune 500 firm's core production process—without its name or industry. Using our model, they could devise an optimal AI-human strategy, whether for Coca-Cola or Tesla, of course adjusting for the sector-specific demands that would guide the specific *type* of AI.

Second, firms with highly modular, decomposable tasks should prioritize sophisticated AI over human capital. Here, AI can dominate core process workflows while humans oversee peripheral administrative decisions. However, such high decomposability is rare. More common are sequenced tasks where AI and humans iteratively refine each other's outputs.

In AI-to-H tasks, managers must develop technical skills to interpret and refine AI's search using context-specific heuristics—avoiding overuse, which can introduce bias. In H-to-AI tasks, performance is maximized through two strategies: (a) supplementing expert human search with sophisticated AI, or (b) when resources are scarce, using even a less sophisticated, hallucinatory AI to explore strategic options that are in the human's blind spot. However, broader applications of hallucinatory AI raise critical discussions on AI ethics, trust, and regulation – an area ripe for future research. Table 1 summarizes these managerial implications across task structures and model parameters.

---Insert Table 1 here---

5.2. Which is a better sequence: AI-to-H or H-to-AI?

The discussion above assumes an organization's task structure is fixed or given. However, our findings reveal a deeper insight: managers can *choose* the task sequence—AI first or human first—based on our model, depending on the quality of human capital in the organization. A simple comparison in Figure 6 shows that an H-to-AI sequence, where H is a domain expert, outperforms an AI-to-H sequence. In technical terms, *average APO [AI| expert H] > max APO [H|AI]*. In essence, collaboration is most effective when AI enhances the prior work of a high-performing domain expert.

--Insert Fig. 6 here--

Figure 7 provides a statistical generalization of this powerful insight highlighting three cases of relevance.

--Insert Fig. 7 here--

Case 1 (AI-first, Expert Human second): AI, with its vast data processing power, samples the full distribution of strategic choices and their payoffs. A high-capability human follows, but their domain expertise leads them to focus on the right tail of AI's distribution. This means much of AI's initial search is wasted, lowering the overall process payoff despite the human's expertise. Our simulations vividly illustrate the concept of "AI wastage" in search (Figure 8). We plot the final pay-off against the number of local peaks discovered by the AI in Step 1 and the ones discovered by H in Step 2. Even within the optimal pay-off zone ($4 < N_{AI}/N_{H} < 6$), AI uncovers over 3.5 times more local peaks than the human ultimately leverages. The difference between what AI discovers and what H utilizes reflects the extent of "AI wastage"—the surplus insights generated by AI that remain untapped in the human's subsequent adaptive search. This phenomenon resembles an attention-allocation failure (Ocasio, 1997), where the bounded rationality of H prevents the full utilization of AI's expansive search capacity.

--Insert Fig. 8 here--

Case 2 (Expert Human-first, AI second): The expert human, leveraging their domain knowledge, starts with sampling an above-average subset of strategic choices. A sophisticated rule-based AI then refines and extends this search, unlocking even better solutions rather than exploring the entire decision space from scratch. This minimizes AI wastage and maximizes aggregate performance.

Case 3 (Novice Human-first, AI second): A non-expert human, lacking domain knowledge, starts near the left tail of the distribution—essentially getting trapped in a suboptimal peak in their NK landscape. Here, even a hallucinatory AI can improve their choices more effectively than a rule-based AI (which may reinforce their flawed search due to retained memory). However, this path remains less optimal than Cases 1 and 2.

Therefore, all else equal, we posit the following preference order of task sequence in an organization to maximize performance:

Case 2 (AI augments expert human's search) > Case 1 (Expert human refines AI's search) > Case 3 (Hallucinatory AI rescues non-expert human's inferior search).

Emerging real-world scenarios have already highlighted the essence of this finding. Consider, for instance, prompt engineering for an LLM to refine a human's task. It is widely established that a well-formulated prompt from a human expert—"On this data, regress price as a function of demand using demand as a lagged variable, incorporating individual fixed effects"—when augmented by the LLM, yields the highest-quality output (representing Case 2 above). Conversely, asking the LLM for an unformulated menu of options—"Analyze this price-demand data"—and then selecting the best option is suboptimal, reinforcing the limitations of an AI-first approach (Case 1) (Harvard University Information Technology (HUIT), 2023; Acar, 2023).

5.3. Limitations and future research

While our simulation-based model provides foundational insights into AI-human collaboration under varying task structures, there are limitations in our approach that suggest directions for future research. First, our model simplifies the distinct cognitive architectures of humans and AI by representing humans as boundedly rational and heuristic-driven, and AI as either rule-based or hallucinatory. While this dichotomy captures key differences in adaptive behavior, it overlooks more complex, real-world dynamics such as attention-weighted search, reinforcement learning, or hybrid decision-making. Future models could incorporate more granular learning behaviors or dynamic adaptation rules to reflect these realities. Moreover, the stylization of hallucination as memory-less random search, while conceptually useful, may under-represent the structured and semantically rich nature of generative AI "hallucinations." Differentiating types of hallucinations—semantic, structural, or contextual—could lead to more realistic modeling of failure modes.

Second, the simulation abstracts away from the broader organizational structures within which AI-human collaboration unfolds. Factors such as hierarchical role boundaries, coordination costs, and cultural dynamics are not captured. Agent-based extensions that embed AI and humans within layered or team-based organizations could uncover how structural constraints or incentive misalignments affect complementarity. Future research could explore how firms strategically reshape workflows to exploit complementarities—balancing short-term implementation costs with long-term performance benefits.

Third, although our model emphasizes structural and computational factors (search size, interdependence, and sequence), it omits subjective organizational elements such as trust, perceived reliability, ethical guardrails, and interpretability. These elements critically influence whether humans choose to accept, reject, or adjust AI-generated outputs. Future simulation frameworks could endogenize trust as a dynamic variable—affected by prior accuracy, transparency, or explainability—and model how trust modulates the division of labor. Similarly, incorporating ethical constraints or normative biases (e.g., fairness filters or human override thresholds) could improve our understanding of when algorithmic reasoning is socially acceptable versus procedurally risky.

Finally, while our model yields theoretically precise predictions about the optimal sequencing of AI and human search, it remains untested in field contexts. Empirical research—whether through experimental studies, case-based analyses, or real-world implementation data—is essential to calibrate and validate the model's predictions. For instance, studying decision-making in domains such as clinical diagnosis, M&A evaluation, or product development could reveal where our theoretical claims hold—and where additional contingencies (e.g., time pressure, ambiguity tolerance) might moderate the observed effects.

5.4. Conclusion

Daniel Kahneman stated, "You should replace humans by algorithms whenever possible. Even when the algorithm does not do very well, humans do so poorly and are so noisy that, just by removing the noise, you can do better than people (Kahneman, 2018; p. 609)." Our findings partially support the first part of this claim—specifically, for modular tasks, an agentic AI may replace humans unless the human involved is a domain expert engaged in a focused search. However, we demonstrate that the best outcomes arise when AI supplements the prior work of a domain expert human. This is the essence of true AI-human complementarity. Our results challenge Kahneman's assertion that AI should simply take over. Instead, they highlight the importance of re-skilling or up-skilling humans to fully leverage AI's potential for the broader benefit of society. Moreover, we contest the second half of his statement, which categorically frames "noise" as detrimental. In practice, noise—or randomness—often plays a crucial role in problemistic search, helping agents escape local peaks (Rivkin & Siggelkow, 2003; Ganco & Hoetker, 2009). In a related vein, we find that strategically applied noise—such as moderately hallucinatory AI—can enhance performance when introduced after a human, particularly one who is not a domain expert. By unlocking exploration pathways beyond human capability, this form of guided randomness proves beneficial.

Felin and Holweg (2024) also argue that Kahneman's perspective—'AI will and should take over'—is overly extreme, as AI, while inherently strong in predicting based on past data, lacks the "causal reasoning" ability of humans, which is essential for experimentation and the creation of new knowledge (p. 363). Though not explicitly stated, their argument however implies that they view AI as a tool rather than an autonomous agent capable of decision-making through causal reasoning. In our model, this distinction is partly reflected in the contrast between humans' heuristic-based (context-dependent) approach and AI's rule-based (context-independent extrapolation of the past) approach to adaptive search. However, we contend that even if our formulation downplays AI's ability in causal reasoning, AI remains agentic as long as it can adapt its search process based on an independent structure distinct from human training.

Thus, our paper potentially offers a pragmatic resolution to the Kahneman—Felin & Holweg debate; a clash of perspectives that captures the general uncertainty around the topic, 'Can AI do strategic decision-making (with humans)?'. Yes it can, and neither will/should AI fully take over, nor is AI entirely incapable of reasoning like humans as yet. Instead, the AI-human agentic collaboration in organizational decision-making is nuanced: depending on the task structure, and the distinctive approaches to adaptive search by AI and humans, the relative emphasis on one agent over the other, or their joint collaboration, becomes more significant.

In sum, our NK simulation of AI–human joint decision-making yields three key insights with broad organizational relevance. First, by foregrounding task structure as the lens for analyzing AI–human collaboration, we offer a generalizable framework applicable across diverse systems and industries. Second, our finding that the optimal sequence involves an expert human initiating the search, followed by an AI refining those decisions, challenges the prevailing notion that AI should lead as a filter, with humans merely adding post hoc nuance (also see Ai, Langer, Muggleton, & Schmid, 2023) . Third, we show that even a hallucinatory AI can rescue suboptimal human decisions from local optima—highlighting a critical managerial implication: when high-quality human capital is scarce, investing in an imperfect AI may still yield strategic gains by expanding the exploratory potential of decision-making.

REFRERNCES

Acar, O. A. (2023, June 6). AI prompt engineering isn't the future. Harvard Business Review. Retrieved from HBR.

Acemoglu, D., & Restrepo, P. (2018). Artificial intelligence, automation, and work. In *The economics of artificial intelligence: An agenda* (pp. 197-236). University of Chicago Press.

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, *40*(S1), S293-S340.

Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous Intelligence for Complex Goals–A Comprehensive Survey. *IEEE Access*.

Adner, R., Csaszar, F. A., & Zemsky, P. B. (2014). Positioning on a multiattribute landscape. *Management Science*, *60*(11), 2794-2815.

Agarwal, S., & Kapoor, R. (2023). Value creation tradeoff in business ecosystems: Leveraging complementarities while managing interdependencies. *Organization Science, 34(3)*, 1216–1242.

Aggarwal, V. A., & Wu, B. (2015). Organizational constraints to adaptation: Intrafirm asymmetry in the locus of coordination. *Organization Science*, *26*(1), 218-238.

Ai, L., Langer, J., Muggleton, S. H., & Schmid, U. (2023). Explanatory machine learning for sequential human teaching. *Machine Learning*, *112*(10), 3591-3632.

Anadol, R., & Kivrak, P. (2024). Beyond human aesthetics: Machine hallucinations. In *Collaborative intelligence: How humans and AI are transforming our world* (p. 125).

Bader, V., & Kaiser, S. (2019). Algorithmic decision-making? The user interface and its role for human involvement in decisions supported by artificial intelligence. *Organization*, *26*(5), 655-672.

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).

Bell, J. J., Pescher, C., Tellis, G. J., & Füller, J. (2024). Can AI help in ideation? A theory-based model for idea screening in crowdsourcing contests. *Marketing Science*, *43*(1), 54-72.

Biggiero, L. (2016). NK simulation modeling. In *Relational methodologies and epistemology in economics and management sciences* (pp. 61-99). IGI Global Scientific Publishing.

Birhane, A. (2021). The impossibility of automating ambiguity. *Artificial Life, 27(1)*, 44–61.

Burton-Jones, A., Butler, B. S., & Scott, S. V. (2021). NEXT-GENERATION INFORMATION SYSTEMS THEORIZING: A CALL TO ACTION. *MIS quarterly*, *45*(1).

Choi, S., Kang, H., Kim, N., & Kim, J. (2025). How does artificial intelligence improve human decision-making? Evidence from the AI-powered Go program. *Strategic Management Journal*, 1–32.

Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. (2025). Human-AI ensembles: when can they work?. *Journal of Management*, *51*(2), 536-569.

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.

Clement, J., & Puranam, P. (2018). Searching for structure: Formal organization design as a guide to network evolution. *Management Science*, *64*(8), 3879-3895.

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.

del Campo, M., & Leach, N. (Eds.). (2022). *Machine hallucinations: Architecture and artificial intelligence*. John Wiley & Sons.

Fang, C., & Kim, J. H. (2018). The power and limits of modularity: a replication and reconciliation. *Strategic Management Journal*, *39*(9), 2547-2565.

Felin, T., & Holweg, M. (2024). Theory is all you need: AI, human cognition, and causal reasoning. *Strategy Science*, *9*(4), 346-371.

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative ai. *Business & Information Systems Engineering*, *66*(1), 111-126.

Ganco, M., & Hoetker, G. (2009). NK modeling methodology in the strategy literature: Bounded search on a rugged landscape. In *Research methodology in strategy and management* (pp. 237-268). Emerald Group Publishing Limited.

Gavetti, G., Helfat, C. E., & Marengo, L. (2017). Searching, shaping, and the quest for superior performance. *Strategy Science, 2(3)*, 194–209.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making*, *62*, 88.

Gigerenzer, G. (2020). What is bounded rationality?. In *Routledge handbook of bounded rationality* (pp. 55-69). Routledge.

Harvard University Information Technology. (2023, August 30). *Getting started with prompts for text-based Generative AI tools*. Harvard University. Retrieved from HUIT.

Hillebrand, L., Raisch, S., & Schad, J. (2025). Managing with Artificial Intelligence: An Integrative Framework. *Academy of Management Annals*, *19*(1), 343-375.

Kahneman, D. (2018). Comment on" Artificial Intelligence and Behavioral Economics". In *The economics of artificial intelligence: An agenda* (pp. 608-610). University of Chicago Press.

Kapoor, R. (2013). Persistence of integration in the face of specialization: How firms navigated the winds of disintegration and shaped the architecture of the semiconductor industry. *Organization Science, 24(4)*, 1195–1213.

Kapoor, R., & Adner, R. (2012). What firms make vs. what they know: How firms' production and knowledge boundaries affect competitive advantage in the face of technological change. *Organization Science, 23(5)*, 1227–1248.

Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press.

Kauffman, S. A. (1995). *At home in the universe: The search for the laws of self-organization and complexity*. Oxford University Press.

Kim, S., Khoreva, V., & Vaiman, V. (2024). Strategic Human Resource Management in the Era of Algorithmic Technologies: Key Insights and Future Research Agenda. *Human Resource Management*.

Knudsen, T. (2024). Agent-Based Modelling for Strategy.

Knudsen, T., & Srikanth, K. (2014). Coordinated exploration: Organizing joint search by multiple specialists to overcome mutual confusion and joint myopia. *Administrative Science Quarterly*, *59*(3), 409-441.

Krakowski, S., Luger, J., & Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, *44*(6), 1425-1452.

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative science quarterly*, *52*(4), 667-694.

Lee, G., & Nythruva, J. (2022). A Mendelian perspective on strategic management: path-dependence and artificial selection in a search for sustainable energy. *Journal of Organization Design*, *11*(3), 95-105.

Leiblein, M. J., & Macher, J. T. (2009). The problem solving perspective: A strategic approach to understanding environment and organization. In *Economic institutions of strategy* (pp. 97-120). Emerald Group Publishing Limited.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management science*, *43*(7), 934-950.

Levinthal, D. A., & Wu, B. (2024). Resource redeployment and the pursuit of the new best use: Economic logic and organizational challenges. *Strategy Science*, *10*(1), 32-47

Levinthal, D., & March, J. G. (1981). A model of adaptive organizational search. *Journal of economic behavior & organization*, *2*(4), 307-333.

Luger, J., Raisch, S., & Schimmer, M. (2018). Dynamic balancing of exploration and exploitation: The contingent benefits of ambidexterity. *Organization science*, *29*(3), 449-470.

Nickerson, J. A., Silverman, B. S., & Zenger, T. R. (2007). The problem of creating and capturing value. *Strategic Organization*, *5*(3), 211-225.

Nishant, R., Schneckenberg, D., & Ravishankar, M. N. (2024). The formal rationality of artificial intelligence-based algorithms and the problem of bias. *Journal of Information Technology*, *39*(1), 19-40.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic management journal*, *18*(S1), 187-206.

Puranam, P., Stieglitz, N., Osman, M., & Pillutla, M. M. (2015). Modelling bounded rationality in organizations: Progress and prospects. *Academy of Management Annals*, *9*(1), 337-392.

Raisch, S., & Fomina, K. (2023). Combining human and artificial intelligence: Hybrid problem-solving in organizations. *Academy of Management Review*, (ja), amr-2021.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of management review*, *46*(1), 192-210.

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management science*, *49*(3), 290-311.

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Artificial intelligence hallucinations. *Critical Care, 27(1)*, 180.

Shick, M., Johnson, N., & Fan, Y. (2024). Artificial intelligence and the end of bounded rationality: a new era in organizational decision making. *Development and Learning in Organizations: An International Journal*, *38*(4), 1-3.

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.

Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, *1*, 25-39.

Suzuki, R., & Arita, T. (2005, September). How niche construction can guide coevolution. In *European Conference on Artificial Life* (pp. 373-382). Berlin, Heidelberg: Springer Berlin Heidelberg.

Vanneste, B. S., & Puranam, P. (2024). Artificial intelligence, trust, and perceptions of agency. *Academy of Management Review.*

Wang, P. (2021). Connecting the parts with the whole: Toward an information ecology theory of digital innovation ecosystems. *MIS quarterly*, *45*(1).

Weiss, G. (Ed.). (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press.

Wheeler, G. (2018). Bounded rationality.

Yayavaram, S., & Ahuja, G. (2008). Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly, 53(2)*, 333–362.

Zenger, T. R. (2002). Crafting internal hybrids: Complementarities, common change initiatives, and the team-based organization. *International Journal of the Economics of Business*, *9*(1), 79-95.

Zhang, Q., Lee, M. L., & Carter, S. (2022, April). You complete me: Human-ai teams and complementary expertise. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1-28).

TABLES & FIGURES

Table 1. Model Summary & Managerial Implications

Task Structure	When AI and H are Complementary	When AI Substitutes for H	Managerial Implications
Modular Tasks ($C = 0$)	• AI search space is moderately broad: $5 < N_{AI}/N_H < 6$ • H has low adaptive sophistication: $K_H/K_{AI} < 0.5$ • OR: H is an expert ($K_H/K_{AI} > 3$), and AI's search space is either focused or too broad	• AI search is too broad and H lacks expertise • H applies excessive heuristics	• Favor AI investment over investment in human capital when task complexity is high • Avoid excessive heuristic adaptation from H
Sequence 1: AI-to-H ($C > 0$)	• AI begins with rule-based search • H applies moderate heuristics: $C/K_{AI} < 0.5$ • Works best when AI space is moderately large: $N_{AI}/N_H = 5$	• H applies excessive heuristics: $C/K_{AI} > 2$ • Especially problematic when AI search space is broad: $N_{AI}/N_H > 5$	• Train managers to interpret and moderately calibrate AI suggestions • Avoid over-dependence on heuristics by H
Sequence 2: H-to-AI ($C > 0$) *A. Rule-Based AI*	• H is high capability, producing quality output in Step 1 • AI performs rule-based refinement on entire H search: $C \geq \lvert N_H \rvert$	• H is low capability and produces poor-quality search in Step 1	• Ensure high-quality H search before deploying rule-based AI • Encourage domain specialization in H
Sequence 2: H-to-AI ($C > 0$) *B. Hallucinatory AI*	• H is low capability with poor Step 1 output • AI decouples from H and explores randomly • Effective even when AI has a very large search space: $N_{AI}/N_H > 15$		• Use AI hallucination strategically to rescue H who lacks domain expertise • Develop and apply ethical and regulatory boundaries to guide exploration, if AI is hallucinatory

Figure 1. Salience of task structures to model agentic AI's collaboration with humans

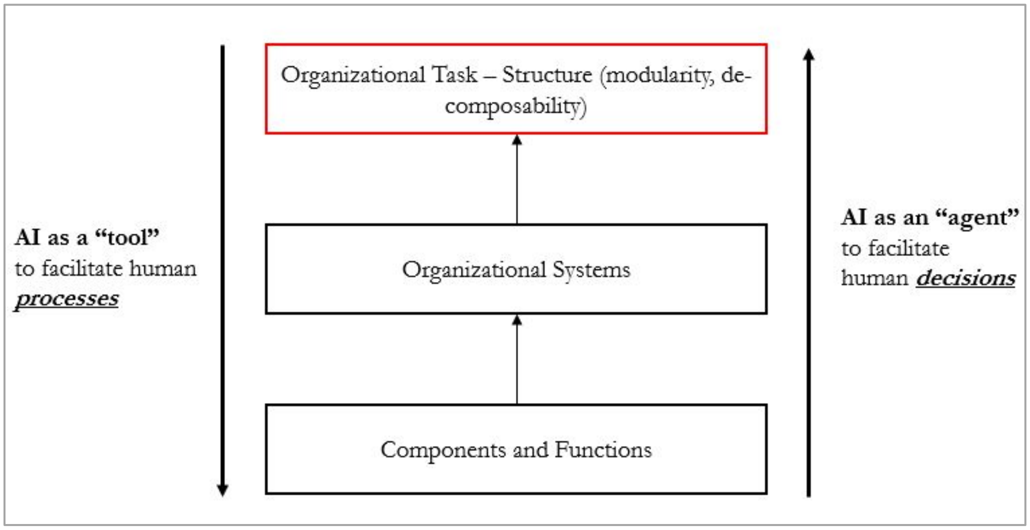
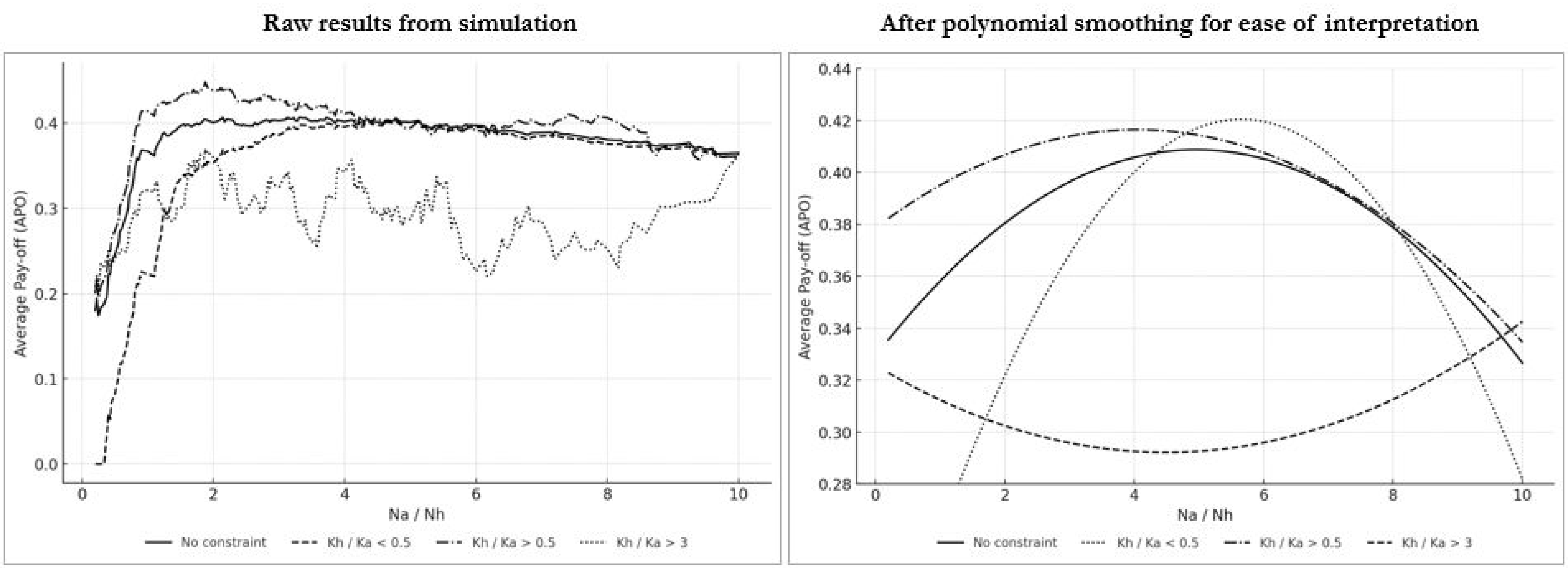


Figure 2. Stylized example of heuristic-based (H) and rule-based (AI) adaptation

Assume: (N = 6, K = 5)	**Heuristic-based: Human**	**Rule-based: AI**	**Hallucinatory: AI**
[0, 0, 0, 1, 1, <u>Pr (N_6 = 1) = ?</u>]			
Prior 5 states in search space, N	[0, 0, 0, 1, 1]	[0, 0, 0, 1, 1]	[0, 0, 0, 1, 1]
Weightages to prior states	[1, 2, 3, 4, 5]	[1, 1, 1, 1, 1]	N/A
Weighted probability calculation of N_6	= [(1x0) + (2x0) + (3x0) + (4x1) + (5x1)]/ [1+2+3+4+5] = 9/15	= 2/5	memory-less coin toss: ~*Bernoulli* (0,1)
<u>Pr (N_6 = 1)</u>	= 0.6	= 0.4	= 0.5

Figure 3. AI-H collaboration in modular tasks (C = 0)

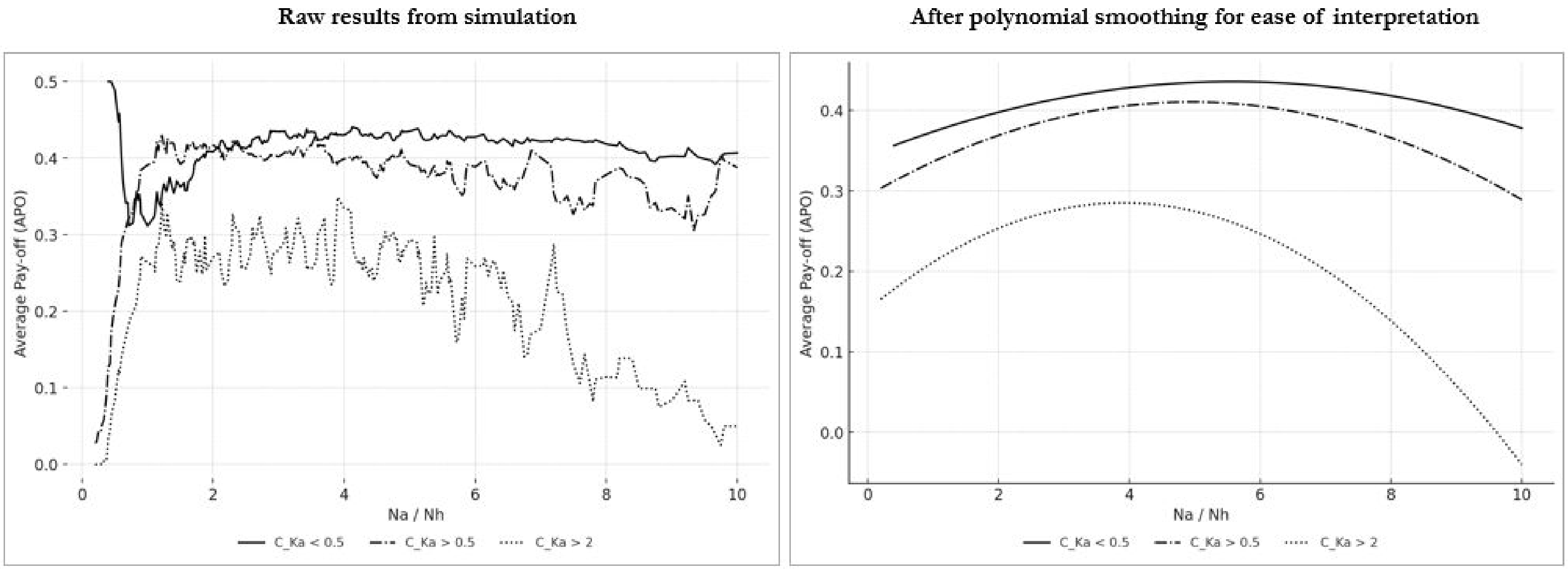


Legend:

- Na/Nh: Ratio of AI to Human search space (always assumed to be > 1)
- Kh/Ka< 0.5: Denotes low human competence/ domain expertise
- Kh/Ka > 0.5: Denotes high human competence/ domain expertise
- Kh/Ka > 3: Denotes extremely high human competence/ domain expertise

Note: Polynomial smoothing done using 'spline' package in Python, allowing polynomial fit up to order 3 (cubic).

Figure 4. AI-H collaboration in sequenced task: AI-to-H (C > 0)

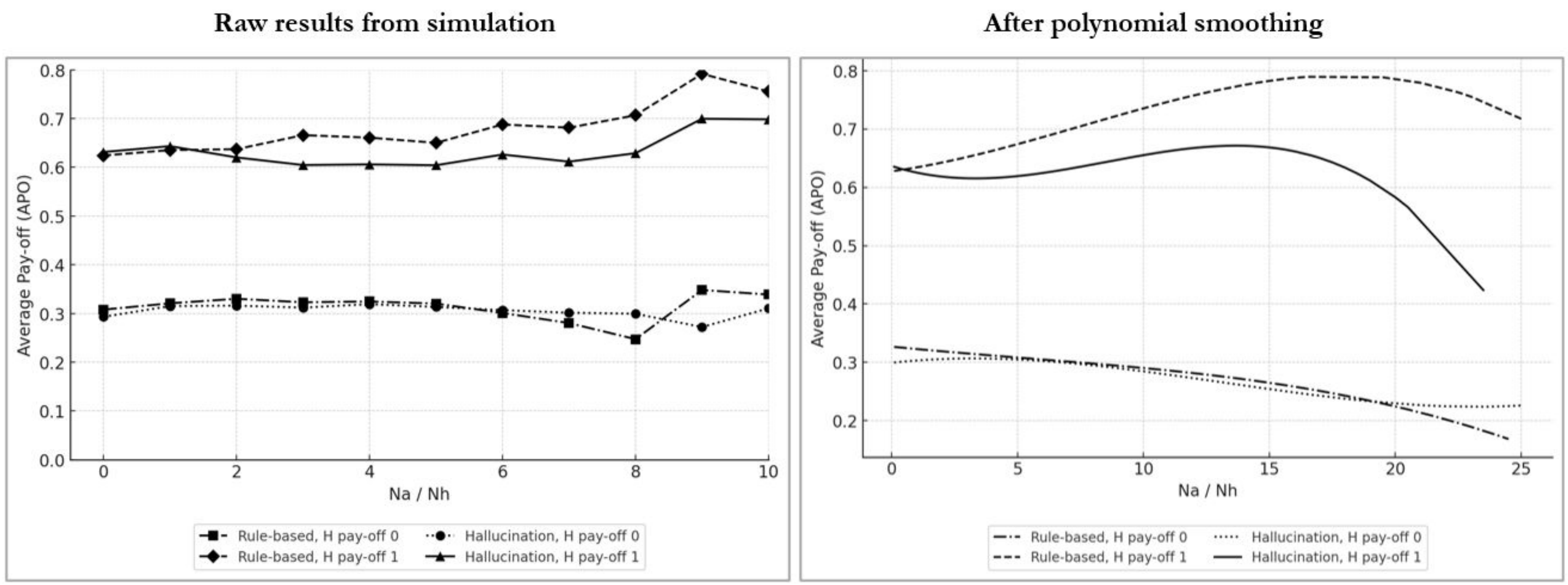


Legend:

- Na/Nh: Ratio of AI to Human search space (always assumed to be > 1)
- C/Ka< 0.5: Denotes mild application of heuristic-based search by H
- C/Ka > 0.5: Denotes strong application of heuristic-based search by H
- C/Ka > 2: Denotes excessive application of heuristic-based search by H

Note: Polynomial smoothing done using 'spline' package in Python, allowing polynomial fit up to order 3 (cubic).

Figure 5. AI-H collaboration in sequenced task: H-to-AI (C > 0)

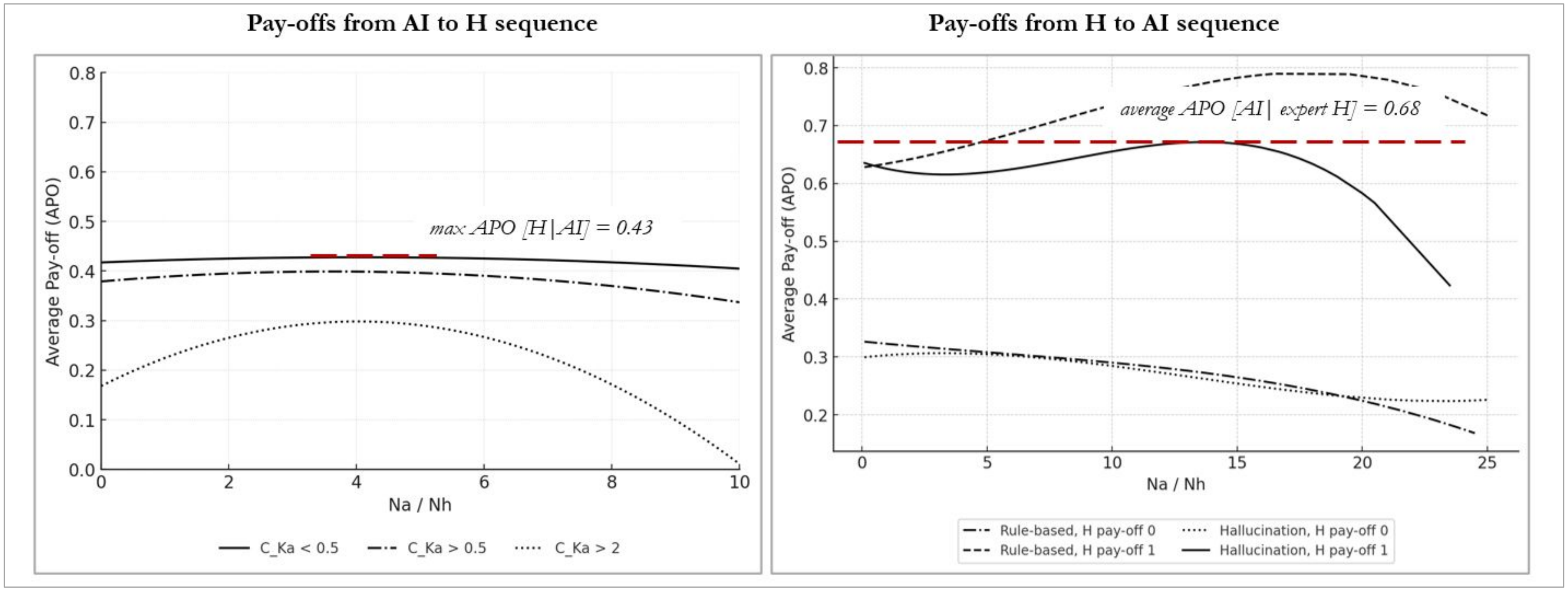


Legend:

- Na/Nh: Ratio of AI to Human search space (always assumed to be > 1)
- H pay-off 1 (high capability H): Intermediate pay-off of H search in step 1 > 0.42 (sample avg.)
- H pay-off 0 (low capability H): Intermediate pay-off of H search in step 1 < 0.42 (sample avg.)

Note: Polynomial smoothing done using 'spline' package in Python, allowing polynomial fit up to order 3 (cubic).

Figure 6. Pay-off Comparison – AI-to-H vs. H-to-AI sequences



Note: Polynomial smoothing done using 'spline' package in Python, allowing polynomial fit up to order 3 (cubic).

Figure 7. Choice of Sequence: Statistical Generalization

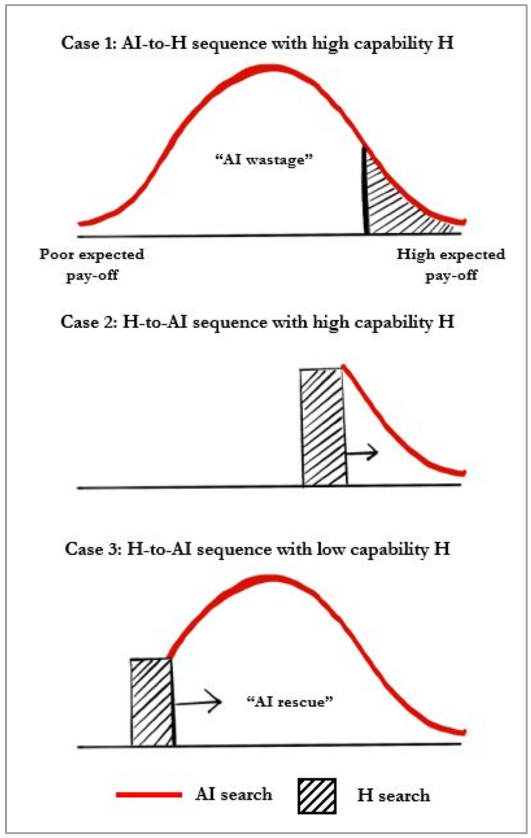
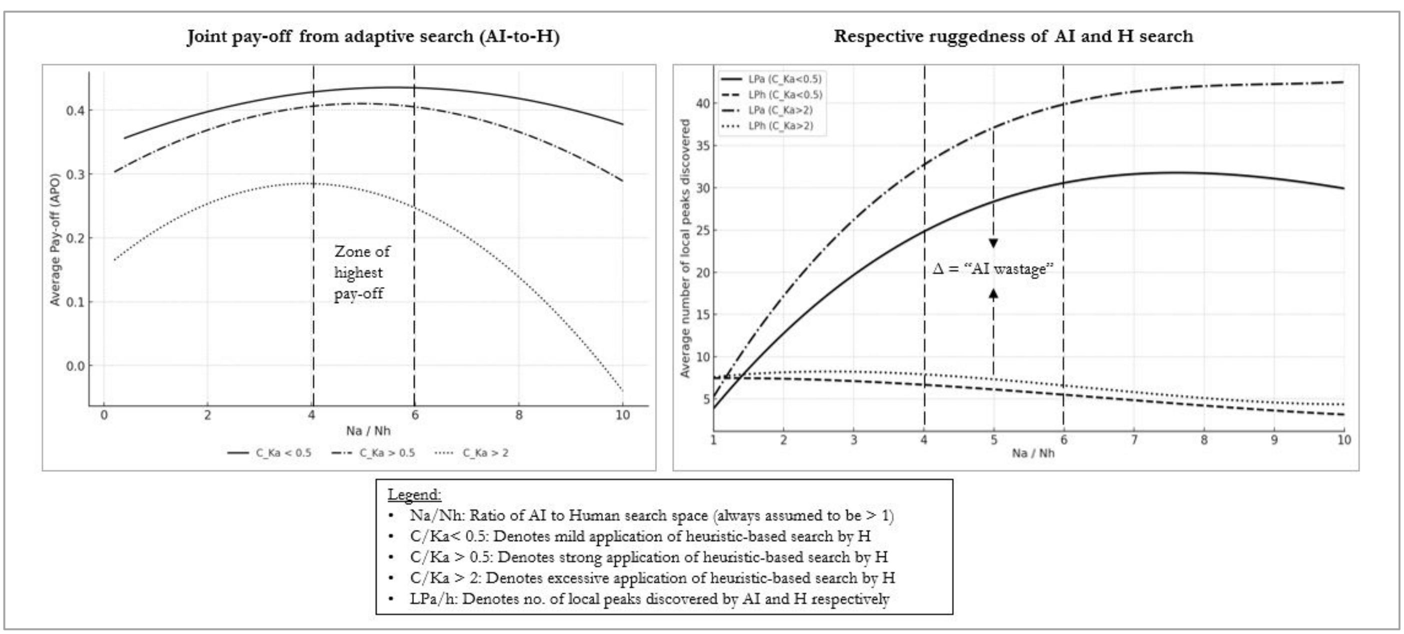


Figure 8. Manifestation of "AI wastage" in our model for AI-to-H sequence



Note: Polynomial smoothing done using 'spline' package in Python, allowing polynomial fit up to order 3 (cubic).