5-11-2023

The Impact of Human-Artificial Intelligence Partnerships on Organizational Learning

Patrick Hendriks
Technical University of Darmstadt, patrick.hendriks@tu-darmstadt.de

Timo Sturm
Technical University of Darmstadt, timo.sturm@tu-darmstadt.de

Christian M. Olt
Technical University of Darmstadt, olt@is.tu-darmstadt.de

Peter Buxmann
Technical University of Darmstadt, buxmann@is.tu-darmstadt.de

THE IMPACT OF HUMAN-ARTIFICIAL INTELLIGENCE PARTNERSHIPS ON ORGANIZATIONAL LEARNING

Research Paper

Patrick Hendriks[1], Technical University of Darmstadt, Germany, hendriks@is.tu-darmstadt.de

Timo Sturm[1], Technical University of Darmstadt, Germany, sturm@is.tu-darmstadt.de

Christian M. Olt, Technical University of Darmstadt, Germany, olt@is.tu-darmstadt.de

Peter Buxmann, Technical University of Darmstadt, Germany, buxmann@is.tu-darmstadt.de

Abstract

To make sense of their increasingly digital and complex environments, organizations strive for a future in which machine learning (ML) systems join humans in collaborative learning partnerships to complement each other's learning capabilities. While these so-called artificial assistants enable their human partners (and vice versa) to gain insights about unique knowledge domains that would otherwise remain hidden from them, they may also disrupt and impede each other's learning. To explore the virtuous and vicious dynamics that affect organizational learning, we conduct a series of agent-based simulations of different learning modes between humans and artificial assistants in an organization. We find that aligning the learning of humans and artificial assistants and allowing them to influence each other's learning processes equally leads to the highest organizational performance.

Keywords: Organizational learning, machine learning, human-AI collaboration, simulation.

1 Introduction

Organizational performance and survival are significantly influenced by organizations' ability to learn about and adapt to the environment (*external reality*) surrounding them (e.g., Levitt and March, 1988; March, 1991; Argote et al., 2003). Since organizations learn through the learning of their members (e.g., Fiol and Lyles, 1985; Levitt and March, 1988; Huber, 1991; Argote et al., 2021), human peculiarities such as their cognitive biases, learning myopia, and bounded rationality have plagued organizational learning from its very beginning (e.g., Simon, 1972, 1979; Levinthal and March, 1993). As learning about the many facets of reality is a challenging endeavor, humans have gradually developed ever more sophisticated tools to enhance their ability to learn (e.g., March, 2006). This has led to the emergence of information systems with ever-increasing information processing capabilities (e.g., Alavi and Leidner, 2001; March, 2006; Berente et al., 2021). Recently, this trend has driven the widespread adoption of artificial intelligence (AI) based on machine learning (ML) algorithms (e.g., Mitchell, 1997; Brynjolfsson and Mitchell, 2017; Russell and Norvig, 2021). By learning from vast amounts of digital information (e.g., digital trace or machine data), ML systems can uncover insights about domains of knowledge that would otherwise remain hidden from humans due to humans' bounded rationality (e.g., Ransbotham et al., 2020). For instance, by better navigating trading markets' vast complexity, ML systems can help identify previously overlooked strategies to improve stock market trading. While non-ML systems only capture human knowledge (e.g., in the form of human-defined rules) and thereby exclusively *support* human learning (e.g., Alavi and Leidner, 2001), ML systems' unique learning

[1] Note: Both authors contributed equally to this paper.

capability allows them to act as a new type of organizational learner and join humans in organizations' search for knowledge with their *own* contributions (e.g., Argote et al., 2021; Sturm et al., 2021).

Since "complexity is all around us in this increasingly digital world" (Benbya et al., 2020, p. 1), we anticipate that the importance of ML systems will continue to increase as more organizational processes rely on digital information and digital domains of knowledge outside of humans' bounded rationality (e.g., Baskerville et al., 2020; Benbya et al., 2020; Kane et al., 2021; Sturm et al., 2021). At the same time, certain domains of knowledge are likely to remain exclusively within the realm of human expertise, as they have yet to be captured or inherently cannot be digitized (e.g., analog or privacy-protected information, intuition, or certain social cues). To complement human learning capabilities, practitioners and researchers alike envision ML-based assistance for all individuals in both private and professional settings (e.g., Kane et al., 2021; Young et al., 2021). For instance, the recent emergence of ML systems that utilize large language models to engage in an ongoing dialogue with individual users increasingly leads to the integration of ML-based assistance for daily work routines (e.g., Microsoft's ChatGPT-based Copilot). Therefore, we conceptualize an *artificial assistant* as an information system that (1) can *learn* and can thus contribute its own knowledge (i.e., is an ML system; e.g., Argote et al., 2021; Sturm et al., 2021), and (2) engages in a *bilateral one-to-one relationship* with a single human partner (i.e., an artificial assistant learns from and adapts to the experiences, suggestions, and feedback provided by an individual and vice versa; e.g., Kane et al., 2021; Young et al., 2021). The emerging interplay, however, requires organizations to effectively coordinate learning between humans, artificial assistants, and each other (e.g., Berente et al., 2021; Young et al., 2021). Due to their complementing learning capabilities, promoting collaborative partnerships between humans and artificial assistants may enhance learning effectiveness (i.e., increase the stock of knowledge), leading to higher organizational performance. Yet, at the same time, poor coordination of such partnerships may cause human or ML-based expertise to be lost or displaced by misinformation, which may threaten organizational survival (e.g., March, 1991; Sturm et al., 2021; Young et al., 2021; Balasubramanian et al., 2022). Thus, we aim to contribute to a better understanding of the learning dynamics between collaborating humans and artificial assistants by asking the following research question: *How can organizations coordinate humans' and artificial assistants' mutual learning to increase organizational learning effectiveness?*

To answer this question, we aim to unpack the dynamics that arise from the interplay between humans and artificial assistants in organizations. Although numerous technologies such as Microsoft's Copilot have been announced recently, the widespread adoption of artificial assistants by organizations is yet to come. As this currently complicates the collection and analysis of empirical evidence, we rely on agent-based simulations to study a future state of an organization in which humans learn side-by-side with artificial assistants (in line with recent long-term-focused research, such as Kane et al., 2021; Sturm et al., 2021; Young et al., 2021). With our explorative study, we hope to preemptively inform managerial strategies that allow an effective future collaboration between humans and artificial assistants.

2 Organizational Learning and Artificial Assistants

To make sense of reality, organizations learn "by encoding inferences from history into routines that guide behavior" (Levitt and March, 1988, p. 319). Organizations thus learn from their past to gradually improve future actions (e.g, Levitt and March, 1988; Argote et al., 2021). As learning thereby essentially defines organizations' actions and resulting performance, organizational learning represents a crucial process that lies at the heart of organizational behavior and organizations' long-term survival (e.g., Huber, 1991; March, 1991; Argote et al., 2021). Although learning is vital to organizations, they cannot learn on their own but rely on their members' learning (e.g., through socialization or education of peers; Fiol and Lyles, 1985; Levitt and March, 1988; March, 1991). Yet, as interactions affect individual learning, organizational learning is more than just the sum of individual learning (e.g., Fiol and Lyles, 1985). This positions organizational learning as a coordination problem, aiming to optimize the learning between organizational members (e.g., Argote and Miron-Spektor, 2011; Argote et al., 2021).

A core issue of organizational learning is the coordination of exploration and exploitation of beliefs (e.g., norms, practices, and routines; March, 1991; Gupta et al., 2006; Argote et al., 2021). While exploration reflects the search for unorthodox beliefs (e.g., experimenting with new technologies or strategies), exploitation represents the reinforcement of established beliefs (e.g., relying on established products and adapting them only gradually; March, 1991; Gupta et al., 2006). As famously demonstrated by March (1991), organizational learning is most effective if it balances exploration and exploitation (e.g., March, 1991; Gupta et al., 2006). If organizations only explore, they will fail to ever make use of developed competence. If organizations only exploit, they will fail to innovate and adapt to pivotal trends in a competitive environment (e.g., March, 1991; Sturm et al., 2021). Yet, achieving the exploration-exploitation balance is challenging. Decades of research have revealed numerous factors that introduce exploitative or explorative tendencies in organizations (excellent overviews exist; e.g., Gupta et al., 2006; Raisch et al., 2009), such as varying organizational structures (e.g., Fang et al., 2010; Schilling and Fang, 2014) and imperfections of learning processes (e.g., Simon, 1991; Levinthal and March, 1993; March, 2010). Especially overcoming the so-called learning myopia (i.e., the tendency to favor exploitation over exploration; Levinthal and March, 1993) still occupies the minds of our greatest scholars (e.g., Levinthal and March, 1993; Balasubramanian et al., 2022; Koçak et al., 2023).

For decades, only humans were able to contribute to organizational learning (e.g., Levitt and March, 1988; Argote et al., 2021). With their ability to learn, ML systems represent a new type of organizational learner contributing their own knowledge (e.g., their own chess, shogi, or Go strategies; Silver et al., 2018) to an organization's stock of knowledge (e.g., Ransbotham et al., 2020; Lyytinen et al., 2021). Surprisingly, research on how information systems affect organizational learning remains in its infancy (e.g., Argote and Miron-Spektor, 2011; Argote et al., 2021). Yet, with the rise of ML, researchers have started to investigate the role and consequences of ML systems in organizational learning (e.g., Argote et al., 2021; Berente et al., 2021). However, research on the impact of ML on organizational learning remains scarce (i.e., Afiouni-Monla, 2019; Seidel et al., 2019; Ransbotham et al., 2020; Lyytinen et al., 2021; Sturm et al., 2021; Balasubramanian et al., 2022), although numerous calls for research strongly stress the lack of research (e.g., Baum and Haveman, 2020; Argote et al., 2021; Berente et al., 2021).

Only recently, scholars have started to discuss consequences of humans and ML systems entering closer relationships (e.g., Schuetz and Venkatesh, 2020; Kane et al., 2021). Although shifts towards closer relationships are known to affect learning (e.g., Lounamaa and March, 1987; Miller et al., 2006; Schilling and Fang, 2014), no study exists that uncovers artificial assistants' impact on organizational learning—despite scholars stressing the great relevance of further unpacking organizational impacts of human-AI relationships (e.g., Grønsund and Aanestad, 2020; Kane et al., 2021; Raisch and Krakowski, 2021). For instance, the integration of ChatGPT with Microsoft Azure may soon enable it to learn from an organization's enterprise data, making ChatGPT's suggestions readily available to individuals when they face work-related issues (Boyd, 2023). So far, however, it is unclear to what extent individuals should rely on the knowledge provided by ChatGPT, verify its suggestions by consulting colleagues, or turn to different knowledge sources—and, in turn, when and how much ChatGPT should calibrate its knowledge to the knowledge of the individuals interacting with it. To address this ambiguity, we now aim to unpack the consequences of learning dynamics between humans and artificial assistants.

3 Simulation Model

We now analyze how the interplay between humans and artificial assistants influences exploration and exploitation in organizations. We first replicate March's (1991) seminal simulation model as our baseline model. Following numerous extensions of March's (1991) model (e.g., Miller et al., 2006; Fang et al., 2010; Butler and Grahovac, 2012; Schilling and Fang, 2014), we then extend the model by adding human interpersonal learning. Finally, we introduce artificial assistants as a new type of learning agent.

3.1 Replicating March's Model of Organizational Learning

March (1991) developed an agent-based simulation model in which an organization's members gradually learn about reality. He considers organizational learning in terms of the mutual learning of the organization and its members. On the one hand, the organization gathers and stores its members' knowledge within norms, practices, and routines, forming the so-called organizational code. On the other hand, each individual within the organization is socialized to the organizational code (i.e., the organization's beliefs) over time. The organizational code thus forms and is formed by individuals in the organization.[2] Following this rationale, March's (1991) model comprises three main entities:

M1: *External reality* is represented as a multidimensional vector consisting of *m* elements, each of which has a value of either 1 or -1. This vector is unknown to the organizational members and independent of their beliefs about it. Initial values are randomly assigned with equal probability.

M2: The organization comprises *n individuals* that each form *m* beliefs about external reality. Beliefs can take on values of either 1 or -1, or individuals can remain agnostic, represented by a value of 0. Initial beliefs are randomly assigned with 1, 0, and -1 having equal probability of occurring.

M3: The *organizational code* reflects the organization's beliefs about external reality and is modeled as an *m*-dimensional vector. Like individuals' beliefs, the organizational code takes values of 1, 0, and -1. The organizational code starts without beliefs about reality and thus initially consists only of zeros.

Individuals and the organizational code can have both correct and incorrect beliefs (i.e., beliefs that do or do not correctly represent external reality). The knowledge levels of all individuals and the organizational code are computed as the percentage of beliefs that match external reality. The average knowledge level represents the average accuracy of beliefs in the organization and is computed as the mean of all individuals' knowledge levels (e.g., March, 1991; Miller et al., 2006). Organizational learning aims to achieve the highest average knowledge level—that is, to maximize the number of matches between individuals' beliefs and the external reality (e.g., March, 1991; Sturm et al., 2021).

Individuals' beliefs and the organizational code change over time (i.e., a sequence of time steps). In each step, individuals learn from the code: For each individual and each of their *m* beliefs, a belief's value changes to the corresponding value in the organizational code with probability p_1. If the organizational code's value is 0, individuals' beliefs are not affected. Learning from the organizational code reflects the socialization of individuals into the organization's beliefs (e.g., norms, practices, and routines). At the same time, the organizational code learns from individuals that are more knowledgeable than itself: For each of the *m* beliefs in the organizational code, the code's value can change to the majority belief held among superior individuals. The probability of learning by the organizational code is a function of the code learning rate p_2 and the level of agreement among the superior individuals (see March, 1991, Footnote 1). Learning by the code represents the adaptation of organizational norms to best practices among the members of the organization (e.g., March, 1991; Miller et al., 2006; Fang et al., 2010).

Based on a series of simulations, March (1991) demonstrated that beliefs in an organization converge over time, eventually leading to a stable knowledge equilibrium. Exploitation occurs when individuals and the organizational code learn from each other rapidly, resulting in premature convergence on homogeneous beliefs that leads the organization to a suboptimal equilibrium. Exploration occurs when slower learning from and by the organizational code conserves belief diversity within the organization. March (1991) observed that slower learning from the organizational code coupled with fast learning by the code lead to the highest average knowledge levels (e.g., Kane and Alavi, 2007; Fang et al., 2010). This means that organizations need to balance exploration and exploitation to optimize their knowledge.

3.2 Introducing Interpersonal Learning and Artificial Assistants

Interpersonal learning. March (1991) did not permit individuals to learn directly from each other, but instead simplified interpersonal learning to learning from and by an organizational code. Decades of research following March's (1991) influential study highlight that much of organizational learning occurs directly from one individual to another in an interpersonal setting and that the emergent dynamics should not be neglected (e.g., Orlikowski, 2002; Borgatti and Foster, 2003; Ethiraj and Levinthal, 2004; Ren and Argote, 2011). Once we consider interpersonal learning, access to potential learning partners

[2] March (1991) does not consider different types of individuals or knowledge to limit the complexity of emerging system dynamics. In line with several other simulation extensions (e.g., Fang et al., 2010; Sturm et al., 2021), we adopt this abstraction as these details are not core to our study objective (following established simulation practices; e.g., Taber and Timpone, 1996).

within an individual's interpersonal network becomes an important factor. Prior research has recognized that people suffer from bounded rationality that causes spatial myopia, which leads them to frequently compare and share their beliefs with a few physically proximate peers (e.g., Simon, 1972; Kauffman and Levin, 1987; Miller et al., 2006). Only if the local search for knowledge among immediate neighbors fails to produce satisfactory results, individuals turn to distant knowledge sources (Cyert and March, 1963). The positive relationship between physical proximity and communication of organizational members is consistent with a large body of research on information flow within organizations (e.g., Monge and Kirste, 1980; Allen, 1984; Zahn, 1991). Though individuals may recognize that some of their peers perform better than themselves (i.e., hold more correct beliefs), they do not know which beliefs in particular lead to better performance (e.g., March, 1991; Kane and Alavi, 2007; Schilling and Fang, 2014). This is corroborated by the fact that an individual's better-performing peers may still have contradicting beliefs about some dimensions of external reality (e.g., how to approach certain routines or decisions). To face this ambiguity, we assume that individuals adopt a majority decision rule similar to the one used to update the organizational code in March's (1991) model: individuals gather the beliefs of other, better-performing individuals with whom they interact, and consider updating each of their beliefs to the majority view. Individuals using a majority decision rule is affirmed by numerous studies on organizational decision making (e.g., Castore et al., 1971; Davis and et al, 1975; Kameda and Davis, 1990). Thus, we follow Fang et al. (2010) and substitute direct interpersonal learning between individuals and their peers for March's (1991) organization-wide learning from the organizational code. In line with prior simulation studies (e.g., Miller et al., 2006; Miller and Lin, 2010), we model interpersonal networks by placing all individuals on a spatial grid with access to distant connections:

E1: The n individuals are situated in a grid in which each individual has four immediate neighbors located to the north, east, south, and west. In order to ensure that all individuals have the same number of neighbors, we constructed the grid so that it has no edges.

E2: Individuals engage in local and distant search for learning partners. In each time step, individuals search locally by identifying better-performing peers (i.e., peers with higher knowledge levels) among their four immediate neighbors. Each individual then determines the majority belief among these better-performing peers for each dimension of the m-dimensional belief vector. For each of the m dimensions, the individual adopts the corresponding majority belief with probability p_{int}. Only nonzero beliefs can be adopted. When there is a tie (i.e., the number of peers with a belief of 1 is equal to that of those with a belief of -1), the individual's belief is not affected. Only if none of their four neighbors perform better than themselves, individuals engage in distant search. Distant search involves identifying the better-performing peers among four remote agents (i.e., all individuals excluding the searcher and their neighbors) that are randomly drawn at the start of the simulation. Individuals again choose to learn from the superior performers using a majority decision rule. This time, if no superior performers are identified, the individual will not learn. In each time step, the n individuals search in random order, since a consistent search order would favor late searchers (e.g., Axelrod, 1997).

Representation of artificial assistants. In March's (1991) model, humans are the only learning agents. However, this assumption is being challenged by the widespread adoption of ML systems. Early ML systems are highly specialized and tailored to specific tasks rather than specific individuals (e.g., Brynjolfsson and Mitchell, 2017; Mitchell et al., 2018; Sturm et al., 2021). However, we recently observe more and more organizations (especially tech giants such as Amazon, Google, and Microsoft) and researchers (e.g., Kane et al., 2021; Young et al., 2021) envisioning personal ML-based assistance for each and every individual in private and professional settings (e.g., OpenAI, 2023). For example, a journalist could team up with a large language model to generate media content. While the first real-world applications of artificial assistants designed to support individuals already exist (e.g., Alexa, Google Duplex, ChatGPT), they do not yet exploit the full potential of ML systems (i.e., they only rarely and limitedly evolve with the assisted individual's needs) as proposed by researchers (e.g., Kane et al., 2021; Young et al., 2021). In line with ML research (e.g., Mitchell, 1997; Russell and Norvig, 2021; Sturm et al., 2021), we assume that artificial assistants rely on ML to derive models that represent reality. March (1991) assumes that all belief dimensions are accessible to all organizational members. However, we can observe that ML systems such as artificial assistants not only contribute knowledge to domains where humans already have some level of expertise, but can also help them gain knowledge about domains that previously were inaccessible to them (e.g., Fügener et al., 2021; Jussupow et al., 2021; Sturm et al., 2021). For example, a broker's artificial assistant could analyze vast amounts of digital trace data to help develop effective trading strategies. However, artificial assistants are no knowledge panacea, as certain knowledge domains have yet to be or inherently cannot be captured in large databases and thus cannot be interpreted by artificial assistants (e.g., Argote et al., 2021). ML systems are already complementing and, in some cases, substituting human decision making (e.g., Jussupow et al., 2021; Raisch and Krakowski, 2021). We anticipate that human dependence on artificial assistants will continue to grow as "reality becomes a reflection of our models in the digital world" (Baskerville et al., 2020, p. 509) and as relevant organizational realities are shifted into digital contexts too complex for humans' bounded rationality (e.g., Baskerville et al., 2020). In line with prior research (e.g., Baskerville et al., 2020; Kane et al., 2021; Sturm et al., 2021), we propose that humans require assistance to comprehend certain dimensions of reality. For this purpose, we introduce artificial assistants as a new type of learning agent that can add to the formation of an organization's norms, practices and routines:

E3: A proportion q_1 ($0 \leq q_1 \leq 1$) of the m beliefs about external reality are interpretable by humans. Another proportion q_2 ($0 \leq q_2 \leq 1$) of the m beliefs about external reality are interpretable by artificial assistants. The two resulting sets of beliefs may intersect (i.e., some beliefs are interpretable by both types of learning agents). We always choose q_1 and q_2 so that the number of human-interpretable dimensions m_{hi} (= q_1m) and artificial-assistant-interpretable dimensions m_{aai} (= q_2m) are integers.

Knowledge base. The diffusion of beliefs (i.e., ML models) among artificial assistants is not tied to geographical proximity or interpersonal networks. ML researchers have proposed federated learning as a technical setup that enables mutual learning between multiple interconnected artificial assistants. In this setup, an integrated ML model is learned through several iterations of updates between a centralized server (i.e., the *knowledge base*) and personalized edge devices that learn local ML models (e.g., Hard et al., 2018; Yang et al., 2019; Ding et al., 2022). Organizations already make use of federated learning to ensure personalization on each local device. For instance, Google uses federated learning for their Android keyboard to personalize writing suggestions locally for one user while gradually integrating all personalized ML models to optimize writing suggestions across all devices (Hard et al., 2018). To capture this notion, we assume that artificial assistants learn from each other only indirectly by updating their beliefs (i.e., their local ML models) through the knowledge base. The knowledge base serves as a centralized knowledge repository for all artificial assistants maintained either by the organization itself, or an external software provider (e.g., Amazon, Google, OpenAI). Learning from and by the knowledge base is modeled after March's (1991) organizational code. Analogous to the mutual learning process between organizational members and the organizational code, diffusion of beliefs through the knowledge base involves two types of interactions: The knowledge base aggregates beliefs about external reality from a subset of artificial assistants. Contributors are selected based on their performance relative to the knowledge base's performance—that is, only superior performing artificial assistants may contribute. A belief about external reality (i.e., a centralized ML model) is derived based on the majority view among superior performers. The knowledge base then pushes its beliefs to all artificial assistants, who will update their local beliefs. This again reflects real-world applications of federated learning (e.g., Hard et al., 2018; Yang et al., 2019; Ding et al., 2022). To incorporate mutual learning between artificial assistants and the knowledge base into our model, we implement the following model extensions:

E4: The knowledge base reflects artificial assistants' aggregated beliefs about reality and is modeled as an m-dimensional vector. Like all learning agents' beliefs, the knowledge base takes values of 1, 0, and -1. The knowledge base starts without beliefs about reality and initially consists only of zeros.

E5: Artificial assistants' beliefs and the knowledge base change over a sequence of time steps. Learning from and by the knowledge base closely follows March's (1991) implementation of learning from and by the organizational code. For each artificial assistant and each of its m beliefs, the belief's value changes to the corresponding (nonzero) value in the knowledge base according to the artificial assistants' update rate p_{aa}. The knowledge base learns by aggregating beliefs from better-performing artificial assistants. Knowledge levels for both the knowledge base and artificial assistants are computed across their interpretable dimensions (i.e., not considering exclusively human-interpretable dimensions). The probability of learning by the knowledge base is a function of the knowledge base's update rate p_{kb} and the level of agreement among superior artificial assistants.

Synchronization. Up to this point, we described two separate models of organizational learning: one model that involves humans and one that involves artificial assistants. However, humans and artificial assistants do not simply co-exist in isolated ecosystems within organizations, but are closely connected to each other. Organizational members in various domains (e.g., human resources, product innovation, sales) have started to collaborate closely with ML systems on a variety of tasks (e.g., Brynjolfsson and Mitchell, 2017; Raisch and Krakowski, 2021). Recently, we can observe that this relationship between humans and ML systems becomes increasingly personalized, evolving from centralized ML systems that are managed and evaluated by data scientists on behalf of the entire organization to the integration of ML systems at the team and individual level (e.g., Fügener et al., 2021, 2022). Artificial assistants can help their human partner explore belief dimensions that were previously inaccessible to them (e.g., Kane et al., 2021; Young et al., 2021). At the same time, artificial assistants can learn from their human partner, for example by using their domain knowledge as the basis for supervised learning (Raisch and Krakowski, 2021). We define synchronization as the process of mutual learning during which humans' beliefs both shape and are shaped by artificial assistants' beliefs. Through synchronization, humans' and artificial assistants' beliefs may become so tightly interwoven that their shared beliefs more accurately represent external reality than either of them could achieve on their own. This includes humans adopting beliefs about dimensions that can only be interpreted by artificial assistants, and vice versa. For example, a broker may not be able to interpret the vast amounts of data that led their artificial assistant to recommend a new strategy to sell a certain stock, but still choose to follow its recommendation (i.e., adopt and act upon one of the artificial assistants' beliefs). Thus, we implement the following extensions:

E6: Each of the *n* humans is assigned *one* artificial assistant. Both humans and artificial assistants form beliefs about external reality. Beliefs can take on values of either 1 or -1, or agents can remain indecisive represented by a value of 0. For both humans and artificial assistants, initial beliefs on their respective interpretable dimensions are randomly assigned with 1, 0, and -1 having equal probability of occurring. Beliefs on non-interpretable dimensions are initialized with all zeros.

E7: In each time step, humans and their artificial assistants engage in synchronization. For each of their *m* beliefs, each human chooses to adopt the corresponding belief of their artificial assistant with probability p_{sync1}. Likewise, for each of its *m* beliefs, each artificial assistant chooses to adopt the corresponding belief of its human partner with probability p_{sync2}. Only nonzero beliefs can be adopted this way. Both directions of synchronization are independent of the learning agents' performance.

Summary. Figure 1 depicts an exemplary organization with four humans and four artificial assistants. Some beliefs about reality can only be interpreted by either humans or artificial assistants; others can be interpreted by both. After their initial interpretation, all learning agents can adopt all beliefs through all learning processes. In each time step, humans first engage in interpersonal learning (p_{int}) with either local or distant peers and then in synchronization (p_{sync1}) with their artificial assistant. Artificial assistants first update their beliefs (p_{aa}) and then synchronize with their human partner (p_{sync2}). After all individuals acted in random order, the knowledge base learns from superior artificial assistants (p_{kb}).

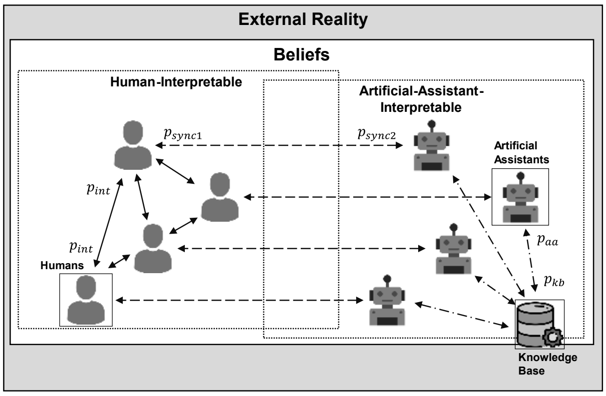

Figure 1. Schematic model representation of organizational learning with artificial assistants.

4 Simulation Results

We constructed our simulation model in Python using the Mesa framework for agent-based modeling.[3] Following previous simulation studies on organizational learning (e.g., March, 1991; Sturm et al., 2021), we tracked the organizational knowledge level (i.e., the average percentage of all learning agent's beliefs that match external reality) for each time step and each simulation run. All simulation runs were terminated after 500 time steps, which we determined sufficient for the organizational knowledge level to converge to its equilibrium value. The reported organizational knowledge levels always refer to the long-term organizational knowledge levels measured at the end of a simulation run. To ensure the robustness of our analysis, the simulation was repeated 100 times for each parameter combination with different random seeds. The results reported in this section are averages calculated from the 100 iterations. All parameters and parameter values used in the simulation are listed in Table 1. To explore learning dynamics that emerge between an organization's human members and artificial assistants, we now analyze the long-term organizational knowledge levels along different parameter combinations.

Parameter	Description	Values
m	Number of external reality and belief dimensions	90
m_{hi}	Number of human-interpretable dimensions	60
m_{aai}	Number of artificial-assistant-interpretable dimensions	60
n	Number of humans (equal to the number of artificial assistants)	50
p_{int}	Interpersonal learning rate: Probability with which humans adopt beliefs from other humans	0.1, 0.3, 0.5, 0.7, 0.9
p_{aa}	Artificial assistants' update rate: Probability with which artificial assistants adopt beliefs from the knowledge base	0.1, 0.3, 0.5, 0.7, 0.9
p_{kb}	Knowledge base's update rate: Determines the probability with which the knowledge base adopts majority beliefs of artificial assistants	0.1, 0.3, 0.5, 0.7, 0.9
p_{sync1}	Humans' synchronization rate: Probability with which humans adopt beliefs from their artificial assistant	0, 0.1, 0.3, 0.5, 0.7, 0.9
p_{sync2}	Artificial assistants' synchronization rate: Probability with which artificial assistants adopt beliefs from their human partner	0, 0.1, 0.3, 0.5, 0.7, 0.9

Note: $p_{sync1} = 0$ and $p_{sync2} = 0$ were included to model an organization without artificial assistants.

Table 1. List of model parameters.

Introducing artificial assistants. We first consider the implications of introducing artificial assistants to an organization. For this purpose, we compare the organizational knowledge levels for different configurations of collaboration between humans and artificial assistants: One configuration in which interpersonal learning is the only source of knowledge for humans as there are no artificial assistants ($p_{sync1} = p_{sync2} = 0$), and other configurations in which humans engage in both interpersonal learning with other humans and synchronization with their artificial assistants ($p_{sync1} > 0$, $p_{sync2} > 0$). Since artificial assistants enable humans to gain insights about belief dimensions that were previously inaccessible to them, introducing artificial assistants inherently increases the number of beliefs that can be formed about external reality (i.e., from 60 to 90 dimensions). To unpack this effect, we now focus on Figure 2 which depicts the organizational knowledge levels calculated based on the *human-interpretable dimensions* (i.e., 60 dimensions; see Figure 2A) and *all dimensions* (i.e., 90 dimensions; see Figure 2B).

Without introducing artificial assistants (i.e., $p_{sync1} = p_{sync2} = 0$; see blue lines in Figures 2A and 2B), the knowledge level in a human-only organization decreases the stronger the humans exploit (i.e., the blue lines decline with increasing interpersonal learning rate p_{int} in Figures 2A and 2B). This detrimental effect of increased human exploitation is consistent with the findings of previous simulation studies that introduced comparable mechanisms of interpersonal learning to March's model and stressed the essential need for human exploration (see, e.g., Fang et al., 2010).

[3] We replicated March's (1991) model by adhering closely to his conceptual description. We validated our replication by qualitatively reproducing the effects of learning rates on organizational knowledge levels (see March, 1991, Figure 1).

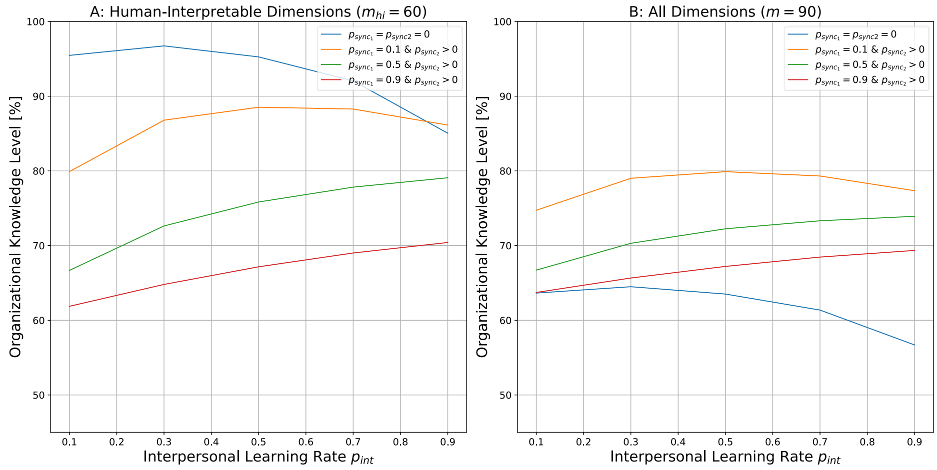

Figure 2. Effect of introducing artificial assistants on organizational knowledge.

When introducing artificial assistants (i.e., $p_{sync1} > 0$ and $p_{sync2} > 0$; see non-blue lines in Figures 2A and 2B), we can observe that the impact of the humans' interpersonal learning on an organization's knowledge level diminishes overall (i.e., the non-blue lines' trend is weaker than the blue lines' trend in Figures 2A and 2B). We can also observe that the trend of humans' interpersonal learning tends to reverse; that is, the stronger humans exploit, the more an organization's knowledge level increases (see the increasing green and red lines in Figures 2A and 2B) or at least remains relatively unaffected (see relatively horizontal yellow lines in Figures 2A and 2B). In line with the findings of Sturm et al. (2021) regarding more generic ML systems that abstract specific human-AI network structures, this observation indicates that an artificial assistant can shift its human partner away from exploration towards a more exploitative role—in other words, an organization using artificial assistants may free its human members from exploration without having to fear significant losses in its long-term knowledge. Thus, we propose:

Proposition 1a: *Artificial assistants can liberate humans from exploration without sacrificing an organization's long-term knowledge.*

When focusing on Figure 2A, we can also observe that the inclusion of artificial assistants leads to a lower knowledge level on the dimensions that are amenable to a human-only setup (i.e., on the 60 human-interpretable dimensions; see the blue line above all non-blue lines in Figure 2A). Thus, at first sight, artificial assistants appear to be detrimental to organizational learning. However, when turning to Figure 2B that captures an organization's knowledge level across all dimensions (i.e., on the 90 human- and artificial-assistant-interpretable dimensions), this presumption no longer holds—if humans and artificial assistants can synchronize (i.e., $p_{sync1} > 0$ and $p_{sync2} > 0$), an organization can reach greater knowledge levels (i.e., all non-blue lines are above the blue line in Figure 2B). While the overall increase in knowledge is not surprising as we add dimensions that cannot be learned in human-only setups, this observation highlights that the coordination of humans and their artificial assistants involves an essential trade-off: While humans and artificial assistants can offer each other access to unique knowledge dimensions, they may simultaneously inhibit each other's learning effectiveness. Thus, we propose:

Proposition 1b: *To reach great long-term knowledge, an organization must coordinate the effects of making further knowledge dimensions available between humans and artificial assistants, and their bilateral influence on each other's learning processes.*

Next, to unpack how an organization can effectively coordinate this trade-off, we analyze the learning dynamics of the two learning subsystems that are included in our simulation and are connected via synchronization: (1) the *learning process of humans* focusing on human-interpretable dimensions and (2) the *learning process of artificial assistants* focusing on artificial-assistant-interpretable dimensions. Remember that Figures 2A and 2B show the *average* of all simulated parameter combinations and thus a mix of all (beneficial and detrimental) coordination strategies for artificial assistants (i.e., p_{aa} and p_{kb}) and the synchronization between humans and artificial assistants (i.e., p_{sync1} and p_{sync2}), leaving it still unclear how organizations can optimize their overall learning. As we have already observed in Figure 2, the existence of artificial assistants can liberate the *human learning process* from exploration, tending to alleviate the importance of humans focusing on more exploration or exploitation (i.e., low or high interpersonal learning rate p_{int}; see Proposition 1a). When turning to the upper quartile of achieved knowledge levels (i.e., the best 25% of configurations), this effect is further pronounced, rendering the specific choice of humans' interpersonal learning rate p_{int} irrelevant. When focusing on the learning process of artificial assistants that resembles March's (1991) original learning system, March's finding of balancing exploration and exploitation proves remarkably robust for artificial assistants despite being connected to the human learning process. Averaged across all configurations and for the upper quartile of achievable knowledge levels, the highest knowledge levels can be achieved with artificial assistants exploring beliefs from their knowledge base and the knowledge base exploiting the artificial assistants' beliefs (i.e., low p_{aa} and high p_{kb}). Those qualitative findings for the two learning subsystems hold true regardless of specific synchronization configurations (i.e., p_{sync1} and p_{sync2}). To find optimal coordination strategies of the interplay of these two subsystems, we do not fixate humans' interpersonal learning (i.e., keep averaging p_{int}) and balance artificial assistants' learning (i.e., $p_{aa} = 0.1$ and $p_{kb} = 0.9$) to capture the best configuration for both subsystems in the following analyses of the subsystems' synchronization.

Synchronization. To effectively manage the observed trade-off between the two subsystems and thereby the collaboration between humans and artificial assistants, we need to understand the subsystems' interplay (i.e., the synchronization rates p_{sync1} and p_{sync2}). For this purpose, we turn to Figures 3A and 3B to examine effects of synchronization between humans and artificial assistants on an organization's knowledge about *human-interpretable* and *artificial-assistant-interpretable* dimensions.

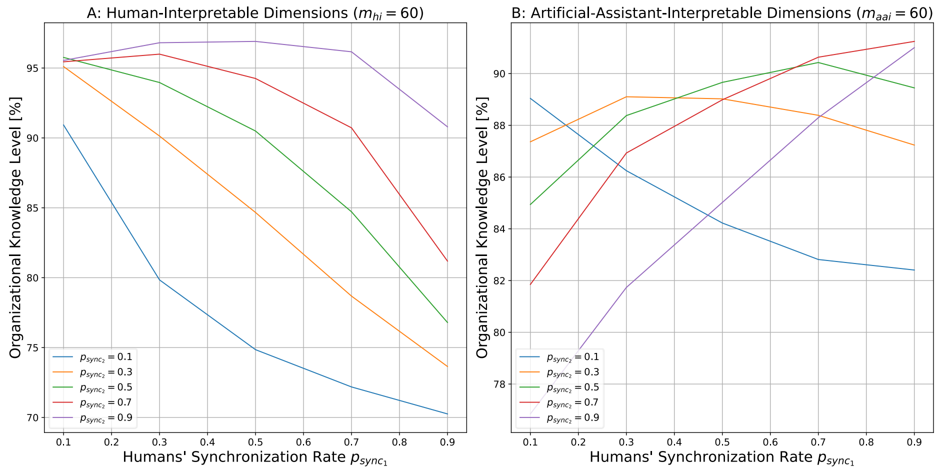

Figure 3. *Effect of synchronization between humans and artificial assistants on organizational knowledge for human-interpretable and artificial-assistant-interpretable dimensions.*

Focusing on optimizing human-interpretable dimensions in Figure 3A, we can observe that the less humans exploit their artificial assistants' beliefs (i.e., decreasing p_{sync1}), the greater an organization's knowledge level regarding human-interpretable dimensions becomes (see decreasing lines in Figure 3A). Yet, the stronger the artificial assistants exploit the beliefs of their human partners (i.e., increasing p_{sync2}), the weaker this effect is pronounced (see the decreasing knowledge gains between high and low p_{sync1} from the blue to the purple line). Thus, an organization can achieve the greatest knowledge levels regarding human-interpretable dimensions when its human members explore their artificial assistants' beliefs while the artificial assistants exploit their human partners' beliefs (i.e., low p_{sync1} and high p_{sync2}). This effect can be explained as follows: Artificial assistants cannot approximate the quality of beliefs about dimensions that are only interpretable by humans, leading to artificial assistants distributing rather random beliefs about human-only dimensions that nullify advances in the human learning process. To circumvent the dissemination of such detrimental random beliefs within the human learning process, reducing the extent of this dissemination benefits achievable knowledge levels regarding human-interpretable dimensions. By coordinating humans to rarely adopt their artificial assistants' beliefs (i.e., low p_{sync1}) and artificial assistants to rapidly adopt human beliefs (i.e., high p_{sync2}), an organization can simply minimize the dissemination of artificial assistants' beliefs within the human learning process. Interestingly, this resembles the tendency of today's organizations who primarily use their ML systems to imitate existing human knowledge (e.g., using ML systems to replicate prior human decisions as closely as possible) while humans remain reluctant to adopt knowledge in domains that are amenable to their own interpretation. Thus, we propose:

Proposition 2a: *To learn human-interpretable dimensions effectively, humans should explore artificial assistants' beliefs while artificial assistants should exploit human beliefs.*

Focusing on optimizing artificial-assistant-interpretable dimensions in Figure 3B, we can observe that the stronger humans exploit the beliefs of their artificial assistants (i.e., increasing p_{sync1}), the greater an organization's knowledge level tends to become (see increasing green, red, and purple lines in Figure 3B). This trend initially weakens and then reverses the less the artificial assistants also exploit the beliefs of their human partners (i.e., decreasing p_{sync2}; see relatively horizontal yellow line and decreasing blue line in Figure 3B). Thus, an organization can achieve the greatest knowledge levels in artificial-assistant-interpretable dimensions when its human members and artificial assistants exploit each other's beliefs (i.e., high p_{sync1} and high p_{sync2}). This can be explained as follows: When humans strongly exploit the beliefs of their artificial assistants (high p_{sync1}), they rapidly adopt beliefs in the short term when artificial assistants' belief diversity is the greatest (March, 1991), which are then strongly exploited by the artificial assistants (high p_{sync2}). As humans cannot approximate the quality of artificial assistants' beliefs, they tend to preserve rather random beliefs of artificial assistants and thereby help prolong belief diversity of artificial assistants' beliefs. In other words, humans serve as an organizational memory of diverse beliefs for artificial assistants. Learning from humans can thereby act as a mechanism that induces exploration into the learning process of artificial assistants (comparable to the effects of a low personnel turnover rate in March (1991), representing a weaker form of employee turnover that has been shown to increase long-term organizational knowledge levels when applied in moderation (March, 1991). Thus, we propose:

Proposition 2b: *To learn artificial assistants-interpretable dimensions effectively, an organization's human members and artificial assistants should exploit each other's beliefs.*

When comparing the optimal coordination of learning about human-interpretable and artificial-assistant-interpretable dimensions (see Propositions 2a and 2b), an essential dilemma arises: while in both cases artificial assistants should exploit human beliefs, humans' optimal learning from artificial assistants' beliefs is contradictory (high vs. low p_{sync1})—So, to learn all dimensions effectively, how should organizations coordinate the synchronization between their human members and artificial assistants? To help unpack this dilemma, we now turn to the Figures 4A and 4B which present an organization's achievable knowledge levels regarding *all* belief dimensions along different synchronization configurations of humans and artificial assistants (i.e., p_{sync1} and p_{sync2}).

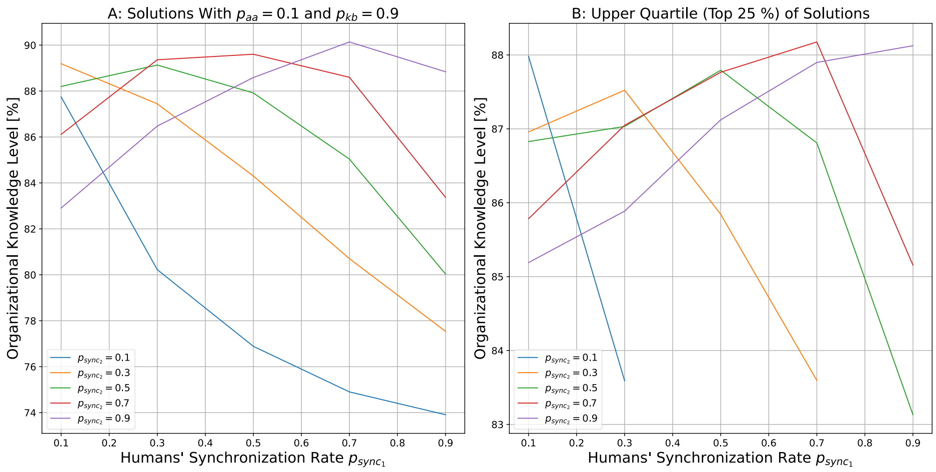

Figure 4. *Effect of synchronization between humans and artificial assistants on organizational knowledge for all dimensions.*

Surprisingly, we can observe in Figure 4A that neither human exploration nor exploitation of artificial assistants' beliefs (i.e., neither high nor low p_{sync1}) generally leads to an organization's greatest knowledge level (see the peaks of the different lines that vary along the x-axis in Figure 4A). Instead, the greatest knowledge levels can be achieved if humans and their artificial assistants tend to learn from each other on equal terms—that is, exploration and exploitation of each other's knowledge are equally valid strategies as long as both humans and artificial assistants follow the same learning strategy (i.e., both tend to explore or exploit equally strong; $p_{sync1} \approx p_{sync2}$). When turning to Figure 4B that shows the upper quartile of achievable knowledge levels (i.e., the best 25% of synchronization configurations), this tendency solidifies and becomes clearly visible (see the peaks of the different lines being perfectly positioned where $p_{sync1} = p_{sync2}$ in Figure 4B). For effective organizational learning of all dimensions, aligning the learning of human members and their artificial assistants therefore appears to represent a worthwhile compromise for organizations aiming to coordinate the involved trade-off between humans' and artificial assistants' contributions. Thus, we propose:

Proposition 3: *To learn all dimensions effectively, an organization's human members and artificial assistants should explore/exploit each other's beliefs with equal intensity.*

5 Discussion

Presently, ML systems increasingly join organizations as a new type of organizational learner, contributing their own knowledge to organizations' stock of knowledge (e.g., Sturm et al., 2021; Balasubramanian et al., 2022). Although research has started to investigate the impact of ML systems on organizational learning (e.g., Seidel et al., 2019; Ransbotham et al., 2020; Sturm et al., 2021), it remains largely unclear how the learning of such human-AI collaborations should be coordinated (e.g., Ransbotham et al., 2020; Argote et al., 2021; Berente et al., 2021). Moreover, research and practice stress the need for closer relationships between humans and ML systems, promoting one-to-one human-AI partnerships with ML systems acting as so-called artificial assistants (e.g., Kane et al., 2021; Young et al., 2021). Although artificial assistants are primarily envisioned to help humans better cope with the complexity of our ever-more digitized world (e.g., Benbya et al., 2021; Kane et al., 2021), the consequences of artificial assistants joining humans in organizational learning remain largely unknown. With our study, we aim to provide new insights into the learning dynamics of humans and artificial assistants and how organizations can effectively coordinate these dynamics.

In particular, we contribute to research that identified the great importance of a deliberate application of artificial assistants (e.g., Kane et al., 2021; Young et al., 2021). Our study confirms the concern that the imprudent use of artificial assistants can, indeed, inhibit human learning processes and thereby jeopardize the loss of human knowledge and organizational competitiveness in the long term. In addition to existing research that highlights the need for an effective technical design of artificial assistants (e.g., Kane et al., 2021; Young et al., 2021), our study stresses that successful use of artificial assistants not only reflects a technical but also a *managerial* issue—that is, managers should expect both potentially virtuous and vicious consequences that are not solely determined by the technical qualities of artificial assistants, but emerge from the specific coordination of artificial assistants' interplay with humans. Our study suggests that focusing solely on optimizing humans' or artificial assistants' learning is not a viable strategy for achieving high levels of organization-wide performance. Instead, the key for increasing organizational learning effectiveness seems to lie in managing the synchronization between humans and artificial assistants: An organization's overall learning requires *both* types of learners to provide each other access to learnings from their unique knowledge domains, which requires the organization to compromise between optimizing both types of learners' knowledge domains.

To address this dilemma, our study suggests that organizations should aim to match the learning modes of humans and artificial assistants (i.e., both exploring or both exploiting each other) to improve organizational learning as a whole by influencing each other's learning processes on an *equal* footing. This requires humans and artificial assistants to be equally receptive to each other's beliefs in order to mutually inform and be informed. Organizations should therefore be cautious about fostering only one type of learner (e.g., through increased automation of human tasks or avoidance of human feedback), and instead encourage the development of mutually rewarding human-AI partnerships.

Recognizing equalized synchronization as a desirable managerial practice is, however, only half the story of managing learning in the era of artificial assistants. Now it becomes crucial to understand how organizations can effectively achieve such an equality. This requires us to identify relevant factors that influence this equality as well as managerial mechanisms that allow organizations to steer humans and artificial assistants towards mutual exploration or exploitation of each other, thereby enabling organizations' control of the synchronization equality. While there are decades of organizational learning research that has examined numerous factors and managerial mechanisms for human exploration and exploitation processes (e.g., Gupta et al., 2006; Argote et al., 2021), the novel artificial assistant context requires us to revisit existing theories and to potentially create new ones to address the peculiarities and unique issues of collaborative human-AI partnerships, which merits further research:

As organizational learning is primarily coordinated by humans, facilitating an equalized synchronization between humans and artificial assistants can be challenging: Knowledge domains that remain exclusively interpretable by artificial assistants due to humans' bounded rationality (e.g., Simon, 1979; Raisch and Krakowski, 2021) limit managers' ability to assess and control the optimization across humans' and artificial assistants' knowledge domains. Hence, the coordination appears to be inevitably biased towards human-interpretable domains of knowledge. This induces the risk of organizational learning favoring the success of human learning at the expense of artificial assistants' knowledge, thereby impairing overall organizational learning effectiveness. Indeed, this human-biased optimization can already be observed with today's organizations using ML systems to exclusively imitate historic human behavior while neglecting to promote potentially useful but unorthodox ML-based ideas. Especially with the advent of large language models (e.g., ChatGPT) that are designed to generate persuasive but not necessarily well-founded statements, determining whether contributions by ML systems represent established knowledge, fruitful original ideas, or harmful misinformation can pose a major difficulty for humans within organizational learning endeavors. In addition, learner-driven issues such as human attitudes toward artificial assistants' beliefs may further complicate managing this synchronization equality. Here, recent phenomena such as algorithm aversion (e.g., Jussupow et al., 2021) or humans' blind obedience to convincing yet false claims made by ML systems (e.g., ChatGPT; OpenAI, 2023) highlight the multifaceted challenges inherent to managing synchronization equality. Especially the latter may also point to novel forms of learning myopia (i.e., the tendency to favor exploitation over exploration; Levinthal and March, 1993) that arise from and reinforce flawed synchronization setups, even though our findings have demonstrated that effective synchronization can also help alleviate existing human learning myopia (supporting research highlighting that effectively coordinated ML systems can help alleviate myopia; e.g., Sturm et al., 2021).

Of course, our study has limitations. By using a simulation, we risk modeling relevant contextual factors and processes as overly simplistic or idealized (e.g., humans are inherently motivated to learn and share their knowledge with others, learn only from knowledgeable peers, and are unaffected by potential power dynamics and cultural differences within an organization). While we tried to reduce this risk by grounding considered factors and processes in a seminal simulation model and comprehensive theory, further research is required to uncover effects of additional factors and process nuances. While our explorative study allowed us to identify potential consequences of abstract organizational learning processes, our findings would greatly benefit from empirical validation and nuance along varying technical and organizational contexts.

We hope that our study can provide a theoretical foundation and stimulate future research on relevant factors and managerial strategies to help advance our understanding of how organizations can optimize learning across the knowledge domains of humans and artificial assistants. While our study is a first step, further research is needed to unpack the emergent dynamics of human-AI partnerships.

Acknowledgment. This research was funded by the German Federal Ministry of Labor and Social Affairs and the European Union through the European Social Fund Plus (ESF Plus) as part of their joint support for the "Future Centers" program.

References

Afiouni-Monla, R. (2019). ‘Organizational learning in the rise of machine learning’, in: *40th International Conference on Information Systems, ICIS 2019*.

Alavi, M. and Leidner, D.E. (2001). ‘Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues’, *MIS Quarterly: Management Information Systems* 25 (1), pp. 107–136. URL: https://doi.org/10.2307/3250961.

Allen, T.J. (1984). *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization*. Cambridge: MIT Press.

Argote, L., Lee, S. and Park, J. (2021). ‘Organizational learning processes and outcomes: Major findings and future research directions’, *Management Science* 67 (9), pp. 5399–5429. URL: https://doi.org/10.1287/mnsc.2020.3693.

Argote, L., McEvily, B. and Reagans, R. (2003). ‘Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes’, *Management Science* 49 (4), pp. 571–582. URL: https://doi.org/10.1287/mnsc.49.4.571.14424.

Argote, L. and Miron-Spektor, E. (2011). ‘Organizational Learning: From Experience to Knowledge’, *Organization Science* 22 (5), pp. 1123–1137.

Axelrod, R. (1997). ‘The dissemination of culture: A model with local convergence and global polarization’, *Journal of Conflict Resolution* 41 (2), pp. 203–226. URL: https://doi.org/10.1177/0022002797041002001.

Balasubramanian, N., Ye, Y. and Xu, M. (2022). ‘Substituting Human Decision-Making With Machine Learning: Implications for Organizational Learning’, *Academy of Management Review* 47 (3), pp. 448–465. URL: https://doi.org/10.5465/amr.2019.0470.

Baskerville, R.L., Myers, M.D. and Yoo, Y. (2020). ‘Digital first: The ontological reversal and new challenges for information systems research’, *MIS Quarterly: Management Information Systems* 44 (2), pp. 509–523. URL: https://doi.org/10.25300/MISQ/2020/14418.

Baum, J.A.C. and Haveman, H.A. (2020). ‘Editors’ comments: The future of organizational theory’, *Academy of Management Review* 45 (2), pp. 268–272. URL: https://doi.org/10.5465/amr.2020.0030.

Benbya, H., Nan, N., Tanriverdi, H. and Yoo, Y. (2020). ‘Complexity and information systems research in the emerging digital world’, *MIS Quarterly: Management Information Systems* 44 (1), pp. 1–17. URL: https://doi.org/10.25300/MISQ/2020/13304.

Benbya, H., Pachidi, S. and Jarvenpaa, S.L. (2021). ‘Special issue editorial: Artificial intelligence in organizations: Implications for information systems research’, *Journal of the Association for Information Systems* 22 (2), pp. 281–303. URL: https://doi.org/10.17705/1jais.00662.

Berente, N., Gu, B., Recker, J. and Santhanam, R. (2021). ‘Managing AI’, *MIS Quarterly* 45 (3), pp. 1433–1450. URL: https://doi.org/10.25300/MISQ/2021/16274.

Borgatti, S.P. and Foster, P.C. (2003). ‘The network paradigm in organizational research: A review and typology’, *Journal of Management* 29 (6), pp. 991–1013. URL: https://doi.org/10.1016/S0149-2063_03_00087-4.

Boyd, E. (2023). *ChatGPT is now available in Azure OpenAI Service*. URL: https://azure.microsoft.com/en-us/blog/chatgpt-is-now-available-in-azure-openai-service/ (visited on 30 March 2023).

Brynjolfsson, E. and Mitchell, T. (2017). ‘What can machine learning do? Workforce implications’ 358 (6370), pp. 1530–1534.

Butler, J.C. and Grahovac, J. (2012). ‘Learning, imitation, and the use of knowledge: A comparison of markets, hierarchies, and teams’, *Organization Science* 23 (5), pp. 1249–1263. URL: https://doi.org/10.1287/orsc.1120.0747.

Castore, C.H., Peterson, K. and Goodrich, T.A. (1971). ‘Risky shift: Social value or social choice? An alternative model’, *Journal of Personality and Social Psychology* 20 (3), pp. 487–494. URL: https://doi.org/10.1037/h0031913.

Cyert, R.M. and March, J.G. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall.

Davis, J.H. and et al (1975). ‘The decision processes of 6- and 12-person mock juries assigned unanimous and two-thirds majority rules', *Journal of Personality and Social Psychology* 32 (1), pp. 1–14. URL: https://doi.org/10.1037/h0076849.

Ding, J., Tramel, E., Sahu, A.K., Wu, S., Avestimehr, S., Zhang, T. and Ai, A. (2022). 'Federated Learning Challenges and Opportunities: An Outlook', in: *ICASSP 2022*. URL: https://doi.org/10.48550/arXiv.2202.00807.

Ethiraj, S.K. and Levinthal, D. (2004). 'Bounded rationality and the search for organizational architecture: An evolutionary perspective on the design of organizations and their evolvability', *Administrative Science Quarterly* 49 (3), pp. 404–437. URL: https://doi.org/10.2307/4131441.

Fang, C., Lee, J. and Schilling, M.A. (2010). 'Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning', *Organization Science* 21 (3), pp. 625–642. URL: https://doi.org/10.1287/orsc.1090.0468.

Fiol, C.M. and Lyles, M.A. (1985). 'Organizational Learning', *The Academy of Management Review* 10 (4), pp. 803–813. URL: https://doi.org/10.2307/258048.

Fügener, A., Grahl, J., Gupta, A. and Ketter, W. (2021). 'Will humans-in-the-loop become borgs? merits and pitfalls of working with AI', *MIS Quarterly: Management Information Systems* 45 (3), pp. 1527–1556. URL: https://doi.org/10.25300/MISQ/2021/16553.

Fügener, A., Grahl, J., Gupta, A. and Ketter, W. (2022). 'Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation', *Information Systems Research* 33 (2), pp. 678–696. URL: https://doi.org/10.1287/isre.2021.1079.

Grønsund, T. and Aanestad, M. (2020). 'Augmenting the algorithm: Emerging human-in-the-loop work configurations', *Journal of Strategic Information Systems* 29 (2), pp. 1–16. URL: https://doi.org/10.1016/j.jsis.2020.101614.

Gupta, A.K., Smith, K.G. and Shalley, C.E. (2006). 'The Interplay between Exploration and Exploitation', *Academy of Management Journal* 49 (4), pp. 693–706.

Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C. and Ramage, D. (2018). 'Federated Learning for Mobile Keyboard Prediction'. URL: http://arxiv.org/abs/1811.03604.

Huber, G.P. (1991). 'Organizational Learning: The Contributing Processes and the Literatures', *Organization Science* 2 (1), pp. 88–115. URL: https://doi.org/10.1287/orsc.2.1.88.

Jussupow, E., Spohrer, K., Heinzl, A. and Gawlitza, J. (2021). 'Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence', *Information Systems Research* [Preprint] (September). URL: https://doi.org/10.1287/isre.2020.0980.

Kameda, T. and Davis, J.H. (1990). 'The function of the reference point in individual and group risk decision making', *Organizational Behavior and Human Decision Processes* 46 (1), pp. 55–76. URL: https://doi.org/10.1016/0749-5978(90)90022-2.

Kane, G.C. and Alavi, M. (2007). 'Information technology and organizational learning: An investigation of exploration and exploitation processes', *Organization Science* 18 (5), pp. 796–812. URL: https://doi.org/10.1287/orsc.1070.0286.

Kane, G.C., Young, A.G., Majchrzak, A. and Ransbotham, S. (2021). 'Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants', *MIS Quarterly* 45 (1), pp. 371–396. URL: https://doi.org/10.25300/MISQ/2021/1578.

Kauffman, S. and Levin, S. (1987). 'Towards a General Theory of Adaptive Walks on Rugged Landscapes', *Journal of Theoretical Biology* 128 (1), pp. 11–45.

Koçak, Ö., Levinthal, D.A. and Puranam, P. (2023). 'The Dual Challenge of Search and Coordination for Organizational Adaptation: How Structures of Influence Matter', *Organization Science* 34 (2), pp. 851–869. URL: https://doi.org/10.1287/orsc.2022.1601.

Levinthal, D.A. and March, J.G. (1993). 'The myopia of learning', *Strategic Management Journal* 14 (2), pp. 95–112.

Levitt, B. and March, J.G. (1988). 'Organizational Learning', *Annual Review of Sociology* 14 (1), pp. 319–338. URL: https://doi.org/10.1146/annurev.so.14.080188.001535.

Lounamaa, P.H. and March, J.G. (1987). 'Adaptive Coordination of a Learning Team', *Management Science* 33 (1), pp. 107–123. URL: https://doi.org/10.1287/mnsc.33.1.107.

Lyytinen, K., Nickerson, J. V. and King, J.L. (2021). 'Metahuman systems = humans + machines that learn', *Journal of Information Technology* 36 (4), pp. 427–445. URL: https://doi.org/10.1177/0268396220915917.

March, J.G. (1991). 'Exploration and Exploitation in Organizational Learning', *Organization Science* 2 (1), pp. 71–87. URL: https://doi.org/10.1287/orsc.2.1.71.

March, J.G. (2006). 'Rationality, foolishness, and adaptive intelligence', *Strategic Management Journal* 27 (3), pp. 201–214. URL: https://doi.org/10.1002/smj.515.

March, J.G. (2010). *The Ambiguities of Experience*. Cornell University Press.

Miller, K.D. and Lin, S.J. (2010). 'Different truths in different worlds', *Organization Science* 21 (1), pp. 97–114. URL: https://doi.org/10.1287/orsc.1080.0409.

Miller, K.D., Zhao, M. and Calantone, R.J. (2006). 'Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model', *Academy of Management Journal* 49 (4), pp. 709–722. URL: https://doi.org/10.5465/AMJ.2006.22083027.

Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., Carlson, A., Dalvi, B., Gardner, M., Kisiel, B., Krishnamurthy, J., Lao, N., Mazaitis, K., Mohamed, T., Nakashole, N., Platanios, E., Ritter, A., Samadi, M., Settles, B., Wang, R., Wijaya, D., Gupta, A., Chen, X., Saparov, A., Greaves, M. and Welling, J. (2018). 'Never-ending learning', *Communications of the ACM* 61 (5), pp. 103–115. URL: https://doi.org/10.1145/3191513.

Mitchell, T.M. (1997). *Machine learning*. New York: McGraw-Hill.

Monge, P.R. and Kirste, K.K. (1980). 'Measuring Proximity in Human Organization', *Social Psychology Quarterly* 43 (1), pp. 110–115. URL: https://doi.org/10.2307/3033753.

OpenAI (2023). *GPT-4 Technical Report*. URL: http://arxiv.org/abs/2303.08774.

Orlikowski, W.J. (2002). 'Knowing in practice: Enacting a collective capability in distributed organizing', *Organization Science* 13 (3), pp. 249–273. URL: https://doi.org/10.1287/orsc.13.3.249.2776.

Raisch, S., Birkinshaw, J., Probst, G. and Tushman, M.L. (2009). 'Organizational ambidexterity: Balancing exploitation and exploration for sustained performance', *Organization Science* 20 (4), pp. 685–695. URL: https://doi.org/10.1287/orsc.1090.0428.

Raisch, S. and Krakowski, S. (2021). 'Artificial Intelligence and Management: The Automation–Augmentation Paradox', *Academy of Management Review* 46 (1), pp. 192–210. URL: https://doi.org/10.5465/amr.2018.0072.

Ransbotham, S., Khodabandeh, S., Kiron, D., Candelon, F., Chu, M. and Lafountain, B. (2020). *Expanding AI's Impact With Organizational Learning*, *MIT Sloan Management Review and Boston Consulting Group*. URL: https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/.

Ren, Y. and Argote, L. (2011). 'Transactive memory systems 1985-2010: An integrative framework of key dimensions, antecedents, and consequences', *Academy of Management Annals* 5 (1), pp. 189–229. URL: https://doi.org/10.1080/19416520.2011.590300.

Russell, S. and Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. 4th edn. London: Pearson.

Schilling, M.A. and Fang, C. (2014). 'When hubs forget, lie, and play favorites: Interpersonal network structure, information distortion, and organizational learning', *Strategic Management Journal* 35 (7), pp. 974–994. URL: https://doi.org/10.1002/smj.2142.

Schuetz, S. and Venkatesh, V. (2020). 'Research perspectives: The rise of human machines: How cognitive computing systems challenge assumptions of user-system interaction', *Journal of the Association for Information Systems* 21 (2), pp. 460–482. URL: https://doi.org/10.17705/1jais.00608.

Seidel, S., Berente, N., Lindberg, A., Lyytinen, K. and Nickerson, J. V. (2019). 'Autonomous tools and design: A triple-loop approach to human-machine learning', *Communications of the ACM* 62 (1), pp. 50–57. URL: https://doi.org/10.1145/3210753.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K. and Hassabis, D. (2018). 'A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play', *Science* 362 (6419), pp. 1140–1144. URL: https://doi.org/10.1126/science.aar6404.

Simon, B.H.A. (1979). 'Rational Decision Making in Business Organizations', *The American Economic Review* 69 (4), pp. 493–513.
Simon, H.A. (1972). 'Theories of Bounded Rationality', *Decision and Organization* pp. 161–176.
Simon, H.A. (1991). 'Bounded Rationality and Organizational Learning', *Organization Science* 2 (1), pp. 125–134. URL: https://doi.org/10.1287/orsc.2.1.125.
Sturm, T., Gerlach, J.P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., Nan, N. and Buxmann, P. (2021). 'Coordinating Human and Machine Learning for Effective Organizational Learning', *MIS Quarterly* 45 (3), pp. 1–22. URL: https://doi.org/10.25300/MISQ/2021/16543.
Taber, C. and Timpone, R. (1996). *Computational Modeling*. SAGE Publications, Inc. URL: https://doi.org/10.4135/9781412983716.
Yang, Q., Liu, Y., Chen, T. and Tong, Y. (2019). 'Federated Machine Learning: Concept and Applications', *ACM Transactions on Intelligent Systems and Technology* 10 (2), pp. 1–19. URL: https://doi.org/10.1145/3298981.
Young, A.G., Majchrzak, A. and Kane, G.C. (2021). 'Organizing workers and machine learning tools for a less oppressive workplace', *International Journal of Information Management* 59 (April). URL: https://doi.org/10.1016/j.ijinfomgt.2021.102353.
Zahn, G.L. (1991). 'Face-to-Face Communication in an Office Setting', *Communication Research* 18 (6), pp. 737–754. URL: https://doi.org/10.1177/009365091018006002.