Predicting the Performance of Human-Agent Collaboration: Insights from Uncertainty Reduction Theory, Dynamic Capabilities Perspectives, and Human Brains

Sohvi Heaton[1]; Jin Ho Yun[2]; Stefan K. Ehrlich[3]

[1] Department of Entrepreneurship and Corporate Innovation, Hankamer School of Business, Baylor University, TX 76706, United States
[2] Wharton Neuroscience Initiative, Wharton School of Business, University of Pennsylvania, PA 19104 United States
[3] SETLabs Research GmbH, Elsenheimerstrasse 55, 80687 München, Germany

ABSTRACT

We investigate the factors influencing the performance of collaboration between humans and autonomous agents, focusing on how and why this performance varies. Our hypotheses are grounded in theories of uncertainty reduction and dynamic capabilities. Results from a laboratory experiment involving approximately 45 minutes of human-agent collaboration per subject indicate that a passive information-seeking strategy affects collaboration performance, particularly when humans observe correct actions performed by the agent. Additionally, people's adaptability, assessed through a measure of dynamic capabilities, positively influences performance. This effect is particularly strong when individuals feel a high level of safety with the agent. Using a multidisciplinary approach, we highlight unique challenges in human-robot interaction, particularly increased uncertainty, to enhance our understanding of how these factors affect the effectiveness of existing theories in human-agent collaborative settings.

1. INTRODUCTION

Robots are increasingly becoming an integral part of many workplaces. As robot technology advances, human-robot teams are achieving increasingly complex tasks (e.g., Evans et al., 2017). Over 750,000 robots are collaborating with employees at Amazon to handle highly repetitive tasks, allowing staff to focus on delivering better service to customers. The latest robot, Sequoia, enhances efficiency by helping employees list items for sale on Amazon.com more quickly and process orders faster and more accurately (Amazon, 2023). As interactions with employees grow, ensuring successful collaboration between humans and robots becomes increasingly vital.

To understand and facilitate successful interactions, Uncertainty Reduction Theory (URT) suggests that individuals can employ various strategies to reduce uncertainty. For example, they can use information-seeking strategies to actively gather information about others. This can be achieved through direct questions (interactive strategies) or by observing the other person's behavior (passive strategy).

Applying URT to human-robot interaction presents unique challenges and strategies. While human-human interactions heavily rely on complex social cues that we naturally interpret, human-robot interactions often demand clearer communication and transparency from the robot to minimize uncertainty (Natarajan et al., 2023). This brings attention to the effectiveness of passive information-seeking strategies in this context. Humans excel at understanding intricate social cues and behaviors from one another, which fosters mutual understanding and reduces the need for significant adaptation (Memar, 2018). In contrast, human adaptation in human-robot interactions can be quite different. Individuals may need to learn specific commands or methods to communicate effectively with robots, which may involve adapting to various interfaces or discovering new ways to give instructions. Moreover, humans must exercise greater caution and awareness of safety protocols when interacting with robots, especially in collaborative environments where physical interaction is involved. This differs from human-human interactions, where safety concerns are typically lower, as people are generally aware of each other's presence and actions (Tong et al., 2024).

This raises questions about the applicability of established theories regarding successful team performance in the context of human-robot collaboration. When integrated into human-nonhuman ensembles, AI-related technologies fundamentally change our understanding of how and why routines evolve or maintain stability (Murray et al., 2020). To date, research on AI in management has primarily focused on how organizations adopt and utilize AI technologies (e.g., Alsheibani et al., 2020; Pumpkin et al., 2019). Some scholars have explored the dual applications of AI, investigating whether AI complements or substitutes human capital (e.g., Choudhury et al., 2020; Fountaine et al., 2019; Daugherty & Wilson, 2018). Others have examined the potential impacts of AI on various forms of human capital (Jia et al., 2023). While some researchers have acknowledged the blurred lines between AI's complementary and substitution effects (Raisch & Krakowski, 2023), they have yet to frame this relationship as collaborative.

Despite advancements, significant theoretical and methodological gaps remain. Previous studies often frame AI as either replacing or augmenting human roles, overlooking the nuanced collaboration between humans and AI. This limits our understanding of optimizing human-agent collaboration. Additionally, there is a lack of comprehensive frameworks addressing the dynamic nature of human-AI interactions and the mechanisms facilitating their integration. This dynamism adds uncertainty to the interaction, a characteristic more pronounced in human-agent interactions, similar to the initial meeting of two strangers (Berger & Calabrese, 1975), as individuals often have limited experience with their machine counterparts. Most research focuses on algorithms and data analytics for decision-making and efficiency, but human-agent interaction involves robots with AI engaging with humans in physical environments. AI allows robots to interpret human cues, enabling both physical and cognitive collaboration, which requires additional, unexplored factors for success. For instance, how humans perceive an agent's actions can influence performance, yet this is difficult to measure through post hoc surveys. This gap hinders the development of a process model for human-robot interaction, which could clarify performance variations in collaboration. In this article, we explore the interplay between humans and autonomous agents to address the question: What factors predict the performance of collaboration between humans and autonomous agents?

To investigate this question, we integrate uncertainty reduction theory with the dynamic capabilities framework. The central premise of uncertainty reduction theory (URT) is that people are inherently motivated to reduce the uncertainty that characterizes initial relationships to better predict and explain others' behaviors. Individuals employ information-seeking strategies to mitigate this uncertainty (Berger & Bradac, 1982). Given human-robot interaction often involves greater uncertainty compared to human-human interaction due to differences in communication styles, understanding, and predictability. This uncertainty can heighten the importance of adaptation on the human part to ensure effective and safe collaboration. Human adaptation can be viewed through the lens of the dynamic capabilities framework—a theory that often examines the interaction between adaptation and uncertainty (Shoemaker et al., 2018). The dynamic capabilities framework posits that the value of possessing dynamic capabilities increases as environmental uncertainty rises. By integrating these theories, we expect that the performance of collaboration between humans and autonomous agents is influenced by humans' passive information-seeking strategies (observing the behavior of the agent) and their dynamic capabilities, both of which may differentially affect collaboration performance.

We test our model through two experimental studies involving human subjects, each conducting a joint task with an autonomous agent (embodied as a cursor on a computer screen or a UR10 robotic manipulator)[1]. Using electroencephalogram (EEG) technology, we collect data on the cortical electrical activity of human brains, associated with the passive information-seeking strategies of humans, along with self-reported perceptions of safety towards the agent and behavioral measures of dynamic capabilities (i.e., response times). In doing so, we make several contributions to the scholarship on AI in management and the underlying theories we draw from.

This study contributes to the growing body of research on AI applications in management (e.g., Aron et al., 2011; Brynjolfsson et al., 2019; Sun et al., 2019) by identifying key determinants, including physical and psychological cues, that influence human-agent collaboration. Additionally, we draw theoretical and methodological insights from strategic management, robotics, and neuroscience, addressing calls for interdisciplinary approaches to fully understand AI's impact on management practices (e.g., Venkatasubramanian et al., 2020). In doing so, our study shifts the theoretical discussion toward a dynamic model of human-AI collaboration, which has been largely absent in the literature, partly due to methodological

[1] Partial data is available at https://github.com/stefan-ehrlich/HRC_neurobased_taskplanning and is provided by Ehrlich et al. (2023).

challenges. We propose that how humans perceive the action of the collaborating agent partly explains the performance of the collaboration between humans and agents. As Popper (1957, pp. 39-40) argued, examining the dynamic processes of a phenomenon helps explain "how and why something happens."

Moreover, we enhance the literature by examining the potential impact of individual attributes such as dynamic capabilities. Previous studies often treat the potential variability across individuals as a nuisance or error variance, which can obscure differences between levels of their independent variables at a general level (Vogel & Awh, 2008). However, previous studies suggest that the performance of collaboration between humans and robots can vary significantly based on the individual (Hopko et al., 2022). The literature on robotics suggests that high adaptability and autonomy are essential design features of the controller (Beer et al., 2014; Heerink et al., 2010). The potential for human adaptation in response to technological constraints needs to be better understood (Caleb-Solly et al., 2018). This adaptability can be harnessed to address some of the limitations of robots. High adaptability fosters greater levels of trust due to a stronger perception of the robot's behavior (Fischer et al., 2018), ultimately resulting in successful collaboration.

Finally, our integrative approach enhances the theories from which we draw. Therefore, we reinvigorate the dynamic capabilities framework within these discussions by elucidating how dynamic capabilities can influence collaboration performance with modern intelligent machines. While the dynamic capabilities framework is generally applied to human-to-human interactions (e.g., Grant et al., 2011; Grijalva & Harms, 2014), our focus extends its applicability to interactions with intelligent machines. Additionally, our work has practical implications, emphasizing significant ways in which the integration of intelligent machines into the workplace may impact human-agent collaboration practices.

2. CONCEPTUAL BACKGROUND

We draw on uncertainty reduction theory and dynamic capabilities, both of which provide insights into how individuals manage and mitigate uncertainty. According to uncertainty reduction theory (URT), individuals prioritize alleviating uncertainty during interactions to enhance the predictability of each party's behavior (Berger & Calabrese, 1975). A key aspect of URT is that individuals employ three information-seeking strategies to reduce uncertainty: active, passive, and interactive (Berger & Bradac, 1982). In VUCA (Volatile, Uncertain, Complex, Ambiguous) conditions, researchers emphasize the importance of dynamic capabilities for performance (Schoemaker et al., 2018).

2.1. A passive information-seeking strategy for reducing uncertainty

When people encounter uncertainty about one another, it can motivate them to interact or communicate in order to reduce that uncertainty. According to uncertainty reduction theory (URT), individuals' primary concern during interactions is to alleviate uncertainty and enhance the predictability of each party's behaviors (Berger & Calabrese, 1975). The central tenet of URT is that individuals utilize three information-seeking strategies to reduce uncertainty: active, passive, and interactive (Berger & Bradac, 1982).

The uncertainty reduction process is relevant to our research because workers are essentially trying to establish a working relationship with the agent. The interactions can potentially influence outcomes (Berger, 1986). For instance, employing a passive strategy involves observing the autonomous agent unobtrusively without direct interaction. This passive information-gathering strategy used by humans can be effectively measured and explored within an emerging body of research in robotics and neuroscience (Wiese et al., 2017). This line of inquiry suggests that bodily signals from human partners, as well as their cognitive states, can be monitored in near real-time through electro- or psychophysiological signals such as electroencephalography, electromyography, eye tracking, heart rate, or galvanic skin response, or a combination of these methods (see, for example, DelPreto et al., 2020).

Different types of information that humans learn from observing an agent can significantly impact performance. Individuals can gain insights about a collaborating agent by noting both negative and positive information regarding its actions (e.g., errors and non-errors). Errors are inevitable and frequently occur in autonomous robot systems (Steinbauer, 2013). Research has identified a negative correlation between robot errors and task success (Carlson & Murphy, 2005). Robots that do not make mistakes are rated as significantly more trustworthy than those that do (Salem et al., 2015). A survey and storyboard-based simulation revealed that human reactions to low and high error severity differ notably, with error severity correlated with a loss of trust in the robot (Rossi et al., 2017).

While prior research consistently shows that robot errors negatively affect task performance, the impact of these errors on people's perceptions of the robot—particularly regarding perceived trustworthiness—remains inconclusive (Stiber & Huang, 2020). Some studies have found contradictory results, indicating minor to no statistical significance regarding the negative impact of errors on trust (Flook et al., 2019). In fact, some research suggests that participants preferred the robot more when it made mistakes during interactions compared to when it performed flawlessly (Mirning et al., 2017). This phenomenon is commonly referred to as the pratfall effect—an increase in likability due to errors (Aronson et al., 1966). Therefore, we present the following hypotheses:

H1a: A passive information-seeking strategy for negative information about the agent is more positively correlated with collaboration performance than a passive information-seeking strategy for positive information.

H1b: A passive information-seeking strategy for negative information about the agent is less positively correlated with collaboration performance than a passive information-seeking strategy for positive information.

2.2. Dynamic capabilities for reducing uncertainty

In addition to passive information-seeking strategies, increased uncertainty in collaboration between humans and autonomous agents may also necessitate adaptation on the human side. Human workers may find themselves interacting with zero-history autonomous agents, lacking individualized information about these machines that could help reduce uncertainty. As a result, humans might experience high levels of uncertainty and be motivated to exhibit human adaptability to mitigate this uncertainty.

Researchers in strategic management have suggested that the value of dynamic capabilities increases in tandem with rising environmental uncertainty. Navigating a VUCA world—volatile, uncertain, complex, and ambiguous—requires dynamic capabilities to maintain an organization's agility, commitment, and profitability (Schoemaker et al., 2018). Dynamic capabilities were initially conceptualized at the organizational level, but the concept has since been broadened to encompass individuals' abilities to adapt. To navigate uncertainty, individuals with robust dynamic capabilities can stay agile and ready to respond to environmental changes.

Research on human-robot collaboration emphasizes the importance of integrating humans' adaptive capabilities during physical interactions to ensure successful collaboration between humans and agents (Ajoudani et al., 2018). Using the commercially available Roomba robot, Sung et al. (2010) proposed the Domestic Robot Ecology—a framework for understanding the long-term acceptance of robots in the home. They identified four temporal stages: preadoption, adoption, adaptation, and use/retention. During the adaptation phase, individuals become more willing to learn about the robot's technical limitations and affordances, prompting them to modify their environment and behavior (e.g., picking up clutter and rearranging items) to maximize the benefits observed in the earlier stages. Several empirical studies have explored how humans adapt when collaborating with intelligent agents or robots. These studies primarily focus on human performance and the resulting subtle behavioral changes observed during short experiments (e.g. Nikolaidis et al., 2017).

2.2.1. The Moderating Role of Perceived Safety in Dynamic Capabilities and Uncertainty Reduction

So far, we have proposed neural and behavioral mechanisms to explain the dynamics of human-agent collaboration. However, we have implicitly assumed that humans perceive the agent as safe for interaction, which may not always be true. To refine our theory, we suggest that perceived safety serves as a boundary condition for the proposed relationships.

Due to the unique characteristics of robots—such as physical embodiment, intelligence, and decision-making—researchers are increasingly emphasizing the role of trust in human-robot interactions (e.g., Baker et al., 2018; Natarajan et al., 2020). Perceived safety refers to "the user's perception of the level of danger when interacting with a robot, as well as the user's level of comfort during that interaction" (Bartneck et al., 2009). In contrast to physical safety, which focuses on preventing robots from causing harm (e.g., through collisions), perceived safety highlights the significant influence of individuals' psychological perceptions on the acceptability and adoption of robots (Zacharaki et al., 2020). Users may feel unsafe even in the absence of physical risks, as various factors—such as the context of use, comfort, familiarity, sense of control, and trust—impact the overall user experience with a robot (Akalin et al., 2023).

Perceived safety is essential for long-term interaction, collaboration, and acceptance in human-robot interaction (HRI). For HRI to be deemed acceptable, a robot must refrain from actions that could induce fear, surprise, discomfort, or create an unpleasant social situation for humans, even if those actions do not cause any physical harm (Sisbot et al., 2010). As robots are still relatively new to many people, this unfamiliarity heightens the significance of perceived safety. Consequently, individuals may be more cautious and sensitive to safety cues when interacting with robots compared to human interactions (Akalin et al., 2023).

We hypothesize that perceived safety plays a positive moderating role in the effects of dynamic capabilities on collaboration performance. In human-robot interaction, users may adjust their behavior more significantly based on their perception of safety. For instance, they might maintain a greater physical distance from a robot they perceive as unsafe or avoid certain interactions altogether (Akalin et al., 2023). In human-human interaction, while perceived safety does influence behavior, the adaptations are often less pronounced due to the inherent familiarity with human interactions. When humans feel unsafe, they are less likely to modify their behavior or explore new ways of interacting with the robot. This reluctance can hinder the development or activation of dynamic capabilities, as users may prefer to stick to familiar behaviors rather than pursue more efficient or innovative interactions (Hamad et al., 2024). Therefore, we propose the following hypothesis:

Hypothesis 2: The effect of dynamic capabilities on collaboration performance is positively moderated by the level of perceived safety regarding the agent.

3. METHODS

Participants

Twenty-five healthy participants took part in the two experiments. Data from one participant (s07) had to be excluded due to technical issues during recording leading to missing event information. The remaining 24 participants were equally distributed by gender (12 females, 12 males) and had an average age of 26.5 years, with a standard deviation of 3.67 years. Half of the participants collaborated with the agent embodied in a cursor on a computer screen, and the other half with the agent embodied in an industrial robot manipulator. Since the study focused on single-subject effects, a larger sample size was not necessary. However, a sample size greater than 10 was selected to allow for basic statistical analyses and potential group-level conclusions. All participants had a technical background, mostly at the bachelor's level in fields such as electrical engineering, computer science, or mechatronics. All but one participant was right-handed, and all had normal or corrected-to-normal vision. Each participant received written and verbal instructions about the experimental protocol, provided written informed consent, and was compensated afterward. The studies were approved by the ethics review board of the Technical University of Munich, reference numbers 80/20 S-KH and 769/20 S-KK. We also note that the data used in this study were published previously in Dimova-Edeleva et al. (2022), and Ehrlich et al. (2023).

Measures.

Dynamic capabilities. Numerous definitions of dynamic capabilities exist, with different authors emphasizing various aspects. Teece et al. (1997, p. 516) define dynamic capabilities as “the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments.” In this context, competences relate to “patterns of current practice and learning” (Teece et al. 1997, p. 518), enabling firms to adapt their operations (Helfat & Winter, 2011). Zollo and Winter (2002, p. 340) describe dynamic capabilities as “a learned and stable pattern of collective activity” that systematically modifies operating routines for improved effectiveness, introducing a "modifying" capability as a third type of dynamic capability.

The level of analysis for dynamic capabilities varies. Teece et al. (1997, p. 515) focus on the firm level, emphasizing the capacity to renew competences, while Eisenhardt and Martin (2000, p. 1107) address the process level, highlighting the evolution and recombination of resources for competitive advantage. Zollo and Winter (2002, p. 340) emphasize collective activity patterns and behavioral adjustments through learning, as noted by Helfat et al. (2007, p. 3). As the concept has evolved, the analysis level has shifted from collective to individual or microfoundational levels, focusing on "managerial learning and adaptation" (Kor & Mesko, 2013, p.234), "entrepreneurial managers" (Teece, 2016), and "dynamic managerial capabilities" (Adner & Helfat, 2003) as well as "managerial cognitive capabilities" (Helfat & Peteraf, 2015, p. 832).

While there are some measures of dynamic capabilities (e.g., Drnevich & Kriauciunas 2011; Kump et al., 2019), they are not suitable for laboratory data. Following prior work (Wollersheim & Heimeriks, 2016), we assess dynamic capability based on behavioral data, avoiding biases associated with self-assessments and acknowledging the tacit nature of dynamic capabilities (Dosi et al. 2000).

To develop an empirically grounded understanding of dynamic capabilities, we adopt Zollo and Winter's (2002, p. 340) definition, emphasizing that routines are fundamental building blocks (Salvato & Rerup 2011). Their definition highlights "learning" and "modification," with learning representing insights from experience (Cohen 1991) and modification relating to internal adjustments (Dosi et al. 2000). Zollo and Winter (2002, p. 340) assert that "dynamic capabilities arise from learning" and are systematic methods for modifying routines. In our study, participants adjusted their routines through learning and procedural modifications due to task novelty. For example, as human and agent partners switched roles, participants had to learn new rules and adapt accordingly. A change in average response speed indicates behavioral adaptation. Specifically, faster response times during task completion signify "improved effectiveness" (Zollo & Winter, 2002). We used changes in response time to measure dynamic capability and assess participants' effectiveness.

To address tautology issues (e.g., Helfat & Peteraf, 2009), we focused on "process improvement" rather than "performance" (Helfat et al. 2007, p. 3). Our study design reflects this by relying on behavioral measures, assessing responses to novelty after defining dynamic capabilities as learning and modifying routines (Zollo & Winter 2002). We operationalized dynamic capabilities in terms of speed, reflecting the ability to "play quickly" to complete tasks. Attention checks were incorporated into our experiment, ensuring that our measures accurately reflected effectiveness following environmental changes.

Passive information seeking strategies: Previous studies measuring passive information-seeking strategies (i.e., observation) often rely on surveys asking respondents about their search behaviors. Examples include viewing a Facebook friend's profile, blog posts, or posted photographs without influencing the target or any other Facebook user (Anderson, 2024) and statements like, "If I had concerns about my private health matter, I would seek information from another source rather than my supervisors" (Li & Lee, 2020).

However, these traditional self-report methods can be subjective. Neuroscientific approaches can provide objective data by directly measuring brain activity, reducing potential bias. A more effective method would utilize naturally occurring brain signals during specific observations. One such signal is the error-related potential (ErrP) which we use to measure human observations of the actions of a collaborating autonomous agent. The ErrP is a specific form of event-related potential (ERP) linked to brain processes implicated in error- and performance monitoring, and while being an established marker for neuroscientific studies has also been proposed as a marker for the assessment and adaptation of robot behavior in human-robot interaction (Ferrez & Millan, 2005, Ehrlich 2020, Ehrlich & Cheng 2016).

This ErrP signal is consistently generated when a person consciously or unconsciously recognizes that an error has occurred, even if the error was made by someone else (Dimova-Edeleva et al., 2022; Ferrez & Millan, 2005).

Human-robot collaborative task

The experimental task involves human-agent interaction in a grid-world environment, where the participants work together with an autonomous agent to complete a trajectory-following task. The task takes place in a 7x7 grid, with both the human and the agent collaborating to move the cursor (or robot's end-effector) across specific tiles along a designated path. The objective is to guide the cursor through each tile in the correct sequence to complete the trajectory. A trial is defined as a single movement of the robot’s end-effector from one grid tile to another. To avoid any visual habituation, the start and end points of the trajectory are randomly chosen, ensuring that the path always includes at least one directional change. For the remainder of this paper, the terms agent and robot are used interchangeably.

Two collaboration scenarios are tested: (1) Sequential Collaboration (sC): In this scenario, the workspace is divided into two colored sections (green and blue), with one section assigned to the human and the other to the robot. Each participant is responsible for controlling the end-effector within their designated area, with control switching at the boundary between the two sections. (2) Intermittent Collaboration (iC): Here, both the human and robot have shared responsibility for the entire trajectory but are limited in their control. The human can only move the robot up and left, while the robot controls the down and right movements. Together, they collaborate to control all four directions.

Takeovers occur when control shifts between the human and robot. In the sequential scenario, the takeover happens at the boundary between their respective areas. In the intermittent scenario, takeovers depend on the movement direction. There is no limit to how many takeovers can happen.

Errors in control can occur at any time. The robot has a 30% chance of making an error, and this may or may not result in a human takeover. Corrections could be handled by either the human or the agent, depending on the direction of movement relative to the trajectory. Throughout the task, human and agent roles are clearly defined, allowing participants to anticipate takeover situations, and fostering mental preparation. These anticipation periods, especially just before a takeover, are of particular interest in the analysis of the data. Trials involving these anticipatory moments are critical for understanding the cognitive processes during collaboration.

The experiment consisted of 12 blocks in total, alternating between the sC and iC scenarios, with each block containing 13 episodes. An episode included all the necessary trials—each being a single movement from one tile to a neighboring one—to complete the marked trajectory from the start to the end point. Upon successfully completing an episode, a new episode with a new trajectory began. After the EEG setup was completed, participants first performed a test block for each scenario, consisting of five episodes per scenario, to ensure they fully understood the experimental setup. The actual experiment then commenced, with participants collaborating with the robot and taking short breaks between blocks as needed. Questionnaires were administered before and after the experiment to account for potential outliers based on prior experience with robots.

Experimental setups are illustrated in Fig 1A and 1B. In experiment 1, participants were seated in front of a computer screen displaying the grid environment; the agent embodied as a red cursor moving from tile to tile controlled either by the agent or the subject. In experiment 2, participants were seated in front of a UR10 industrial robot, embodying the agent, whose end-effector was placed above a screen, displaying the grid environment. Further technical details are described in the original papers by Dimova-Edeleva et al. (2022) and Ehrlich et al. (2023).

INSERT FIGURE 1 ABOUT HERE

EEG data processing

EEG data was collected using a Brain Products actiChamp amplifier with 32 active gel-based electrodes, arranged according to the extended international 10–20 system, covering locations such as (FP1, FP2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, C3, C4, T7, T8, CP5, CP6, P3, P4, P7, P8, TP9, TP10, O1, O2, Fz, Cz, Pz, EOG1, EOG2, EOG3). The electrodes were referenced to the average of TP9 and TP10, and the sampling rate was set at 1,000 Hz. Three additional channels captured EOG signals from specific facial points (left and right outer canthi and forehead), following the method suggested by Schloegl et al (2007).

Data preprocessing was performed using functions from the Matlab EEGLAB toolbox (Delorme & Makeig, 2004). First, a zero-phase Hamming-windowed sinc FIR band-pass filter with cutoff frequencies of 1 Hz and 40 Hz was applied to the EEG and EOG signals to eliminate high-frequency and power-line noise. EOG artifacts such as eye blinks and lateral movements were corrected using a regression-based method from Schloegl et al. The artifact-contaminated EEG channels were identified using normalized kurtosis, following which spherical interpolation was used to reconstruct rejected channels from the signal of neighboring electrodes. The data were then down-sampled to 250 Hz to decrease computation time for the subsequent

The remaining artifacts were reduced using independent component analysis (ICA), which was performed using the infomax algorithm (runica). The resulting independent components (ICs) were submitted to the EEGLAB ADJUST plug-in for automatic identification (Mognon et al., 2011) and removal of ICs associated with generic discontinuities, eye movement, facial muscle, and neck tension-related artifacts. Finally, the data were then re-referenced using a common average reference (CAR) to further reduce external noise.

The data from all participants were segmented into epochs by extracting time intervals from -200 to 1000 ms relative to the onset of the cursor/robot movement (with t = 0 ms marking the start of the cursor/robot movement). These epochs were then categorized into five groups: Human non-error/error trials, e.g. correct and incorrect cursor/robot moves which were controlled by the subject, interface error trials, e.g. incorrect cursor/robot moves which were controlled by the subject, but, despite correct subject command wrongly executed (simulating a faulty interface), and agent non-error/error trials, e.g. correct and incorrect cursor/robot moves which were controlled by the autonomous agent. Baseline correction was applied to each trial and channel individually by subtracting the average amplitude of the -200 to 0 ms period from the entire signal epoch.

We then examined the temporal activation patterns, our neurophysiological index; i.e., average amplitudes of the event-related potentials (ERPs) at channel Cz (mid-central area) in pre-defined timeframes generated while participants observed the robot under the various experimental conditions described earlier. Channel and timeframes were chosen according to the ERP components relevant for brain processes linked to error- and performance monitoring, e.g. the event-related negativity (ERN) observed between 80-150 ms post stimulus, the P300 component observed between 250-500 ms post stimulus, and the Late Positive Potential (LPP), observed between 400-800 ms post stimulus.

4. RESULTS

We began by testing hypothesis 1, examining the relationship between specific event-related potential (ERP) components, our neurophysiological index, and joint task performance. The LPP, which occurs between 400-800 ms after stimulus onset, was a significant predictor of collaboration performance during autonomous agent-controlled trials. Specifically, higher LPP activation during agent non-error trials (observations of the agent's correct actions) was associated with a greater percentage of correct responses from humans and improved collaboration performance ($b = 6.545$, $t(23) = 2.775$, $p = 0.011$). This suggests that the cognitive processes reflected in the LPP component are linked to fewer errors, indicating better performance.

Similarly, the P300 component, was also a significant predictor of performance (b = 2.067, $t(23) = 2.133$, $p = 0.044$). The P300 component is typically associated with attention allocation and conflict monitoring. In our study, higher P300 amplitudes during human non-error trials predicted better performance in terms of human error responses (vs. correct responses). This finding indicates that attentional focus and cognitive conflict resolution during human-controlled actions contribute directly to the participants' ability to make errors. Our results on LPP and P300 components support H1a and H1b. Our results on the LPP and P300 components indicate that both H1a (negative information-seeking strategies) and H1b (positive information-seeking strategies) are significant. However, the positive information-seeking strategy (H1b) appears to be a stronger predictor.

We then examined the relationships among human dynamic capabilities, perceived safety, and collaboration performance as proposed in H2. As mentioned earlier, we calculated the change in average response speed, which reflects behavioral adaptation, and used it as a proxy for dynamic capabilities. This variable emerged as a significant predictor of collaboration performance ($b = 62.773$, $t(23) = 2.799$, $p = 0.010$). A steeper slope, indicating slower adaptation and responses, was positively correlated with a higher percentage of human error responses, suggesting decreased collaboration performance. This finding indicates a significant positive main effect of human dynamic capabilities on collaboration performance.

As predicted, a significant moderation effect was observed between human dynamic capability and self-reported safety perception regarding robots ($b = 2.259$, $t(23) = 2.198$, p = 0.039), alongside a main effect of the LPP component during agent-controlled trials ($b = 5.542$, $t(23) = 2.492$, $p = 0.021$), predicting joint task performance. We express the following OLS regression model (LPP of agent-controlled trials: $F(2, 21) = 6.936$, $p = 0.005$; P300 of human-controlled trials: $F(2, 21) = 7.597$, $p = 0.003$):

$$\text{Joint task performance} = \beta_0 + \beta_1 \cdot \texttt{ERP Component (LPP or P300)} + \beta_2 \cdot \texttt{Dynamic Capability} \times \texttt{Perceived Safety} + \epsilon_{ti}$$

The positive relationship between dynamic capabilities and collaboration performance was stronger for subjects who perceived the robot as safer, as indicated by a score of +1 SD in safety perception. In other words, when participants felt safer towards the agents, they were better able to utilize their dynamic capabilities, which consequently improved their joint task performance. The floodlight moderation analysis (Johnson-Neyman technique) further revealed that this interaction was significant when safety perception was within the range of 9.61 to 14.01 ($p < 0.05$; see Figure 2). Outside of this range, the effect of response time slope on performance became non-significant, suggesting that safety perception plays a critical moderating role among those who perceive the agent safer.

INSERT FIGURE 2 ABOUT HERE

Similar to the results of agent-controlled trials, the same relationships are observed in human-controlled trials. A significant moderation effect was also observed between human dynamic capabilities and self-reported safety perception on agents ($b = 5.867$, $t(23) = 2.999$, $p = 0.007$). Additionally, a main effect of the P300 component during human-controlled trials ($b = 1.907$, $t(23) = 2.293$, $p = 0.032$), predicting collaboration performance. Participants who perceived the robot as safer (+1 SD in safety perception) displayed a stronger positive relationship between their response time improvement and performance. The floodlight analysis further supports this interaction: when safety perception was within the range of 9.14 to 14.49 ($p < 0.05$). Taken together, the moderation effects observed in both agent-controlled trials and human-controlled trials support our H2.

These findings underscore the importance of human observations of the collaborating robot's actions, as measured by emotional evaluation (LPP), attentional processing (P300), and behavioral adaptation of human dynamic capabilities. Together, these factors are crucial for optimizing human performance and ultimately enhancing collaboration outcomes. Moreover, perceived safety emerged as a key factor that strengthens the relationship between human dynamic capabilities and improved performance, reinforcing the notion that psychological factors, such as safety perception, can significantly impact human-robot collaboration. The results suggest that not only do the technical aspects of the interaction–often highlighted in previous studies–matter, but also the emotional and cognitive state of the participants, particularly their perception of safety, which can significantly influence their error-correction abilities.

In sum, our analysis provides evidence that cognitive processes, reflected in the LPP and P300 brain components, along with behavioral adaptations of dynamic capabilities, are crucial for improving performance in error-prone and collaborative environments. Higher safety perceptions further amplify these effects, offering important insights for enhancing human-robot interactions in strategic settings.

5. DISCUSSION

We integrated theories of uncertainty reduction and dynamic capabilities to create a new model that explains how and why collaboration performance between humans and autonomous agents varies. In addressing the how and why questions, we demonstrated that uncertainty reduction acts as an intervening mechanism through which passive information-seeking behaviors employed by humans influence performance. To our knowledge, there is no existing research that explicitly tests this effect. Through our conceptual model testing, we answered calls to utilize multiple theories in explaining human-agent interaction phenomena (e.g., Rocha et al., 2023) and provided a stronger theoretical foundation for understanding how and why neural and behavioral factors at the individual level impact human-agent collaboration performance.

We proposed that a reduction in uncertainty regarding the actions of the agent with which humans interact is a key mechanism underlying the relationship between human responses to the agent's behavior and collaboration performance. We found that a passive information-seeking strategy, particularly regarding positive information about the robot, influences collaboration. Several mechanisms may explain this outcome. First, positive reinforcement plays a significant role. When a robot makes a correct move, it enhances the human’s trust and confidence in the robot’s abilities. This positive reinforcement fosters smoother and more efficient collaboration (Ehrlich et al., 2023). Previous works showed in the context of human-agent co-adaptation, that positive reinforcement plays a greater role than negative for successful learning of consensus policies between human and agent (Ehrlich & Cheng 2018, 2019, Ehrlich 2020).

A second factor may involve reduced cognitive load. When robots perform tasks correctly, it alleviates the cognitive burden on humans, as they do not need to intervene or correct the robot. This leads to more seamless and less stressful collaboration (Mutlu et al., 2016). These elements are all essential for effective collaboration.

As predicted in H2, our findings indicate that the effect of human dynamic capabilities is moderated by perceived safety. Perceived safety enables individuals to be more adaptable in their roles, allowing them to adjust their actions in response to the robot's behavior. In essence, since perceived safety is crucial for establishing trust, human dynamic capabilities are more effectively harnessed when there is trust in the robot. Therefore, ensuring both the physical and psychological safety of human collaborators is essential.

Theoretical Implications

Our interdisciplinary approach enhances our understanding of human and robot collaboration, which is inherently dynamic and context-dependent. It also helps us explore the mechanisms behind successful human-agent collaboration, revealing relationships among variables that may be obscured in studies examining AI and human roles in isolation rather than as part of a dynamic and interactive system.(Dogru & Reskin, 2020). While existing literature primarily focuses on how AI transforms production processes (Aron et al., 2011; Brynjolfsson et al., 2019; Meyer et al., 2014), understanding of human-AI collaboration for performance improvement remains limited. Our study addresses this gap by identifying key determinants of human-robot collaboration. We propose reframing AI as more than just a component of traditional production processes and adapting theories of human-human collaboration to the context of human-robot collaboration.

We emphasize the heterogeneity of individual attributes in relation to AI, which enhances the theoretical foundation of a crucial emerging topic: whether AI complements or substitutes human capital within firms (Choudhury et al., 2020; Fountaine et al., 2019). Between these two perspectives, there may be a middle ground where human users can orient the use of autonomous agents to their advantage. This middle ground can be beneficial for modeling human interaction with emerging AI-based machines. By incorporating individual differences to model human behavior, we can effectively illustrate how these dynamics operate in various important contexts (Lubinski, 2000). A deeper insight into which individuals are less likely to harbor negative perceptions of AI will help scholars and organizations better identify employees who can benefit from and complement the deployment of AI technologies. Treating human users as a homogeneous group undermines the potential of AI applications. In contrast, recognizing individual heterogeneity allows scholars and firms to pinpoint subgroups of employees with greater concerns about AI, enabling more targeted interventions. This approach fosters greater complementarity and reduces friction in adopting AI technologies in the workplace (Tong et al., 2021). While previous studies have confirmed the significance of perceived safety in technology adoption (Zhang & Yu, 2020), this research focuses on the mediating effect of perceived safety in the relationship between individual dynamic capabilities and collaboration performance. By conceptualizing individual characteristics, perceived risk, and dynamic capabilities within the URT framework, we provide a clearer theoretical lens for understanding the mechanisms that contribute to successful human-agent collaboration.

Our paper contributes to the literature by contextualizing and extending URT to the human-agent interaction domain. This study is one of the first to employ URT to explain how humans respond to the action of autonomous agents. We contribute to the theory by adding nuance to how individuals gather information. Our findings suggest that positive information about the collaborating agent—rather than negative information—predicts collaboration performance when humans engage in passive information-seeking strategies. This enhances URT by recognizing that the quality and type of information sought can significantly influence the uncertainty reduction process.

By highlighting the role of perceived safety in mediating the effect of humans' dynamic capabilities on collaboration performance, our research identified a crucial psychological factor that influences how dynamic capabilities are utilized. This adds a new dimension to the existing literature, which often focuses more on organizational and strategic aspects. Our finding bridges the gap between psychology and dynamic capabilities, offering a more holistic understanding of how human factors impact the effectiveness of dynamic capabilities in collaborative settings. Together, these findings provide a nuanced understanding of the empirical scope of dynamic capabilities and highlight the need for further research that integrates diverse perspectives (Peteraf et al., 2013) and methods to explore causal relationships. As Whetten (1989: 492) stated, "temporal and contextual factors set the boundaries of generalizability, thereby defining the range of the theory." We hope that our finding can serve as a foundation for future studies to explore other psychological boundary conditions and their impact on dynamic capabilities, further enriching the literature.

In this study, our application of neuroscience offers insights that were previously challenging to uncover through traditional behavioral research in the field of strategic management. By analyzing data from the human brain, we were able to directly quantify the factors influencing collaboration performance, specifically focusing on humans' passive information seeking strategies of observing the action of the interacting robot. Using a sample of

24 participants, we employ EEG to investigate ERP components relevant to error- and performance monitoring to evaluate humans' real time observation of the correct/incorrect action of the interacting agent, a construct that is difficult to measure through conventional self-reported methods. In addition, a tightly controlled research design enabled us to more precisely isolate the nature of the uncertainty reduction mediating mechanism. Our multidisciplinary approach aligns with the diverse range of research methods and perspectives used in strategic management, acknowledging the intricate nature, depth, and complexity of strategic issues (Durand et al., 2017).

Practical implications

Our findings illuminate organizational design decisions, including task allocations (Taylor, 1911) and coordination activities (Thompson, 1967), within the context of both humans and autonomous agents. Understanding how individuals perceive the actions of autonomous agents is essential for their design and regulation. Our findings suggest that when developing robots intended to collaborate with humans, variability in positive behavior should be considered, as humans appear sensitive to such variability. This sensitivity may, in turn, influence the quality of interactions and the acceptance of robots as interaction partners. In addition, recognizing the significance of perceived safety can guide the design of training programs and collaborative systems, optimizing human-robot interactions for improved performance.

CONCLUSION

Our study represents a significant step in bridging neuroscience, strategic management, and robotics to foster scientifically grounded human-robot collaboration. To facilitate successful human-robot interaction, it is crucial to first understand and measure, using neuroscientific methods, how humans respond to robots. In this research, we demonstrate the neuroscientific and psychological mechanisms that can be evoked by a robot. It provides insights into the behaviors that robots should exhibit to elicit specific brain mechanisms. For effective collaboration, robots must activate automatic and often implicit processes within the human brain.

Figure 1. Experimental setup and human-agent collaboration task, with the agent embodied as red cursor on a computer screen, e.g. in study 1 (A), and the agent embodied as industrial robot manipulator, model UR10, e.g. study 2 (B). Illustration of the two collaboration scenarios (C and D). Figure adapted from Figure 1 in Ehrlich et al. 2023.

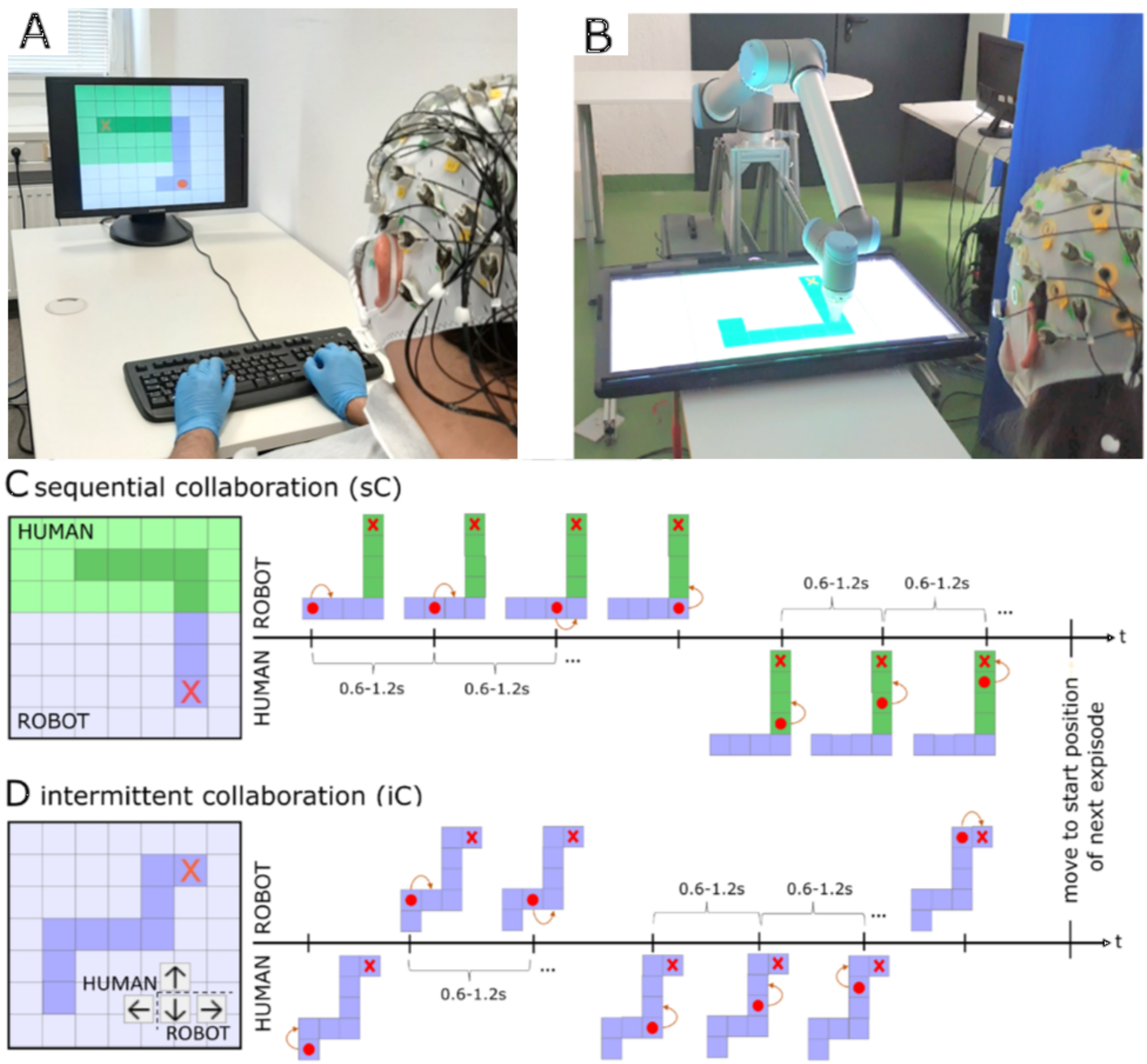
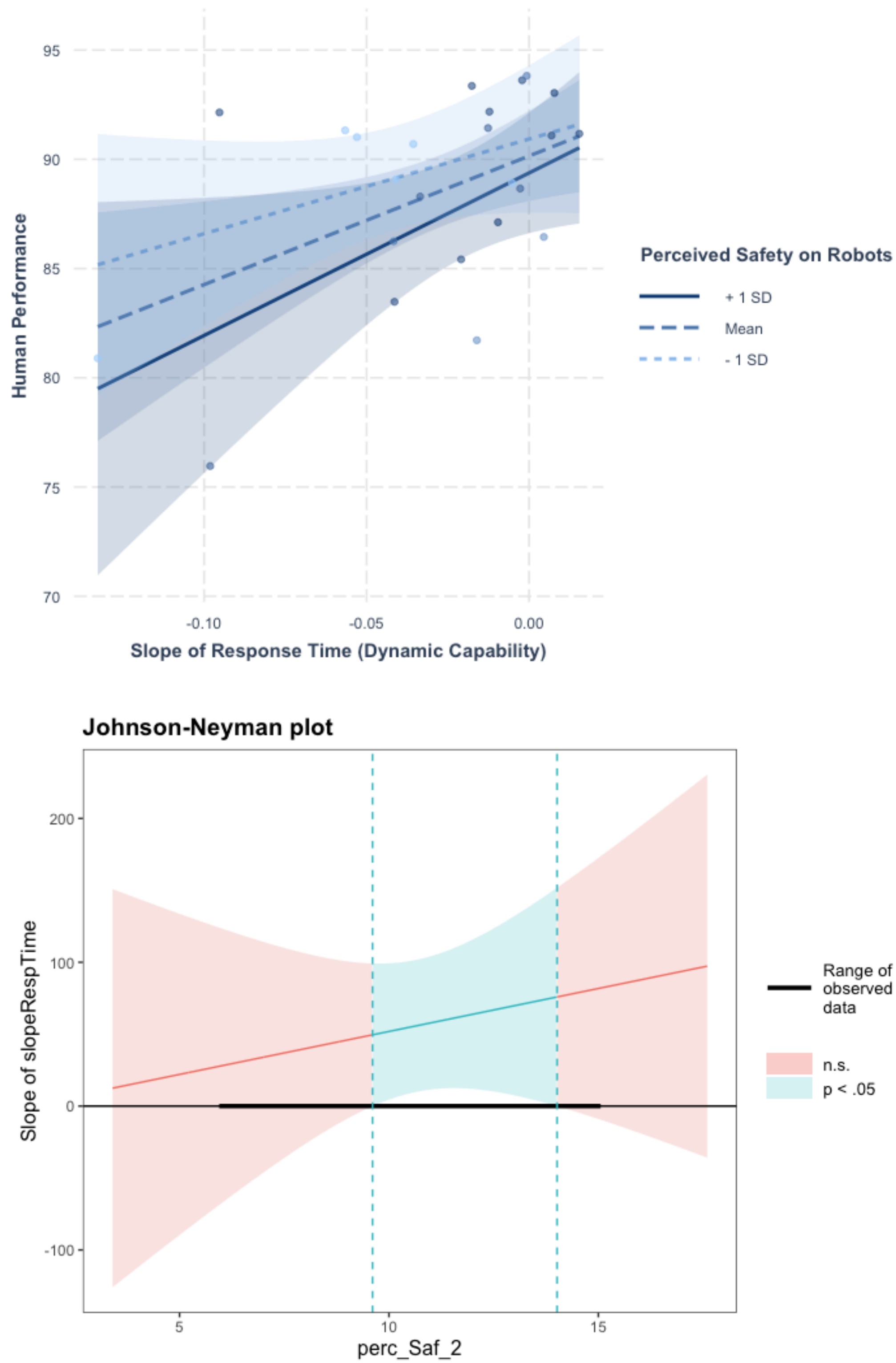


Figure 2. Interaction effects of spotlight and floodlight analyses.



REFERENCES

Adner R., Helfat C. E. (2003). Corporate effects and dynamic managerial capabilities. Strategic Management Journal, 24: 1011-1025.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction machines: The simple economics of artificial intelligence. Cambridge, MA: Harvard Business Press.

Akalin, N., Kiselev, A., Kristoffersson, A., and Loutfi, A. (2023). A taxonomy of factors influencing perceived safety in human–robot interaction. *International Journal of Social Robotics*, pp. 1–12.

Akalin, N., Kristoffersson, A., and Loutfi, A. (2022). Do you feel safe with your robot? Factors influencing perceived safety in human–robot interaction based on subjective and objective measures," *Int. J. Hum.-Comput. Stud.*, 158: 102744

Alimardani M, Hiraki K. (2020). Passive Brain-Computer Interfaces for Enhanced Human-Robot Interaction. Frontiers in Robotics and AI. 7.

Aly, A., & Tapus, A. (2015). Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. *Autonomous Robots*, *40*(2), 193–209.

Alsheibani, S., Cheung, Y., Messom, C., & Alhosni, M. (2020). Winning AI strategy: six-steps to

create value from artificial intelligence. Americas Conference on Information Systems, Online

Amazon. (2023). Amazon announces 2 new ways it's using robots to assist employees and deliver for customers.

https://www.aboutamazon.com/news/operations/amazon-introduces-new-robotics-solutions

Anderson, T. (2024). The use of Facebook for information seeking in social and personal relationships.

Aron, R., Dutta, S., Janakiraman, R., & Pathak, P. A. (2011). The impact of automation of systems on medical errors: Evidence from field research. Information Systems Research, 22(3), 429–446.

Aronson, E., Ben Willerman, andJoanne Floyd.1966.The effect of a pratfallon increasing interpersonal attractiveness. Psychonomic Science 4, 6 (1966), 227–228.

Baker, A., Phillips, E., Ullman, D., Keebler, J. (2018). Toward an understanding of trust repair in human-robot interaction: Current research and future directions. ACM Transactions on Interactive Intelligent Systems (TiiS), 8(4):1-30

Baraglia J, Cakmak M, Nagai Y, Rao RP, Asada M. (2017). Efficient human-robot collaboration: when should a robot take initiative? The International Journal of Robotics Research. 36(5-7):563–579.

Bartneck, C., Kulic, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, 1,: 71–81.

Beer J. M., Fisk A. D., Rogers W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. Journal of Human-Robot Interaction, 3, 74–99

Berger, C. R. & Bradac, J. J., (1982). L*anguage and social knowledge: Uncertainty in interpersonal relations*. London, England: Edward Arnold.

Berger, C. R. & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1, 99-112.

Berger, C. R. (1986). Uncertain outcome values in predicted relationships: Uncertainty reduction theory then and now. *Human Communication Research, 13*, 34-38.

Brougham, D., & Haar, J. 2018. Smart technology, artificial intelligence, robotics, and algorithms (STARA): Employees' perceptions of our future workplace. Journal of Management & Organization, 24: 239–257.

Brynjolfsson, E., Hui, X., & Liu, M. (2019). Does machine translation affect international trade? Evidence from a large digital platform. Management Science, 65(12), 5449–5460.

Brynjolfsson, E., Mitchell, T., & Rock, D. 2018. What can machines learn, and what does it mean for occupations and the economy? American Economic Association Papers and Proceedings, 108: 43–47.

Bughin, J., & Manyika, J. (2019). Your AI efforts Won't succeed unless they benefit employees. Harvard Business Review

Caleb-Solly, P., Dogramadzi, S., Huijnen, C. & Heuvel, H. (2018) Exploiting ability for human adaptation to facilitate improved human-robot interaction and acceptance, The Information Society, 34:3, 153-165

Carlson, J. and Murphy, R. (2005). How UGVs Physically Fail in the Field. IEEE TRANSACTIONS ON ROBOTICS 21, 3 (2005), 423–437.

Chan, T. Y., Li, J., & Pierce, L. (2014). Compensation and peer effects in competing sales teams. Management Science, 60(8), 1965–1984.

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.

Cohen MD (1991) Individual learning and organizational routine: Emerging connections. Organ. Science 2(1):135–139.

Daugherty, P., & Wilson, H. J. (2018). *Human + machine: Reimagining work in the age of AI*. Boston, MA: Harvard Business Review Press.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods, 134(1), 9-21.

DelPreto J, Salazar-Gomez AF, Gil S, Hasani R, Guenther FH, Rus D. (2020). Plug-and-play supervisory control using muscle and brain signals for real-time gesture and error detection. Autonomous Robots. 44(7):1303–1322

Dimova-Edeleva, V., Ehrlich, S., Cheng, G. (2022). Brain computer interface to distinguish between self and other related errors in human agent collaboration. *Scientific Reports*. 12: 20764

Dogru, A. & Keskin, B. (2020) AI in operations management: applications, challenges and opportunities. *Journal of Data, Information and Management*. 2: 67-74

Dosi G, Nelson RR, Winter SG (2000) Introduction: The nature and dynamics of organizational capabilities. Dosi G, Nelson RR, Winter SG, eds. The Nature and Dynamics of Organizational Capabilities (Oxford University Press, New York), 1–22.

Dragan AD, Bauman S, Forlizzi J, Srinivasa SS. (2015) Effects of robot motion on human-robot collaboration. In: 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE: 51–58.

Drnevich PL, Kriauciunas AP (2011) Clarifying the conditions and limits of the contributions of ordinary and dynamic capabilities to relative firm performance. *Strategic Management Journal*. 32(3):254–279.

Durand, R., Grant, R., Madsen, T. (2017). The expanding domain of strategic management research and the quest for integration. *Strategic Management Journal*. 38: 4-16.

Ehrlich, S., Dean-Leon, E., Tacca, N., Armleder, S., Dimona-Edeleva, V., Cheng, G. (2023). Human-robot collaborative task planning using anticipatory brain responses. PLos One.

Ehrlich, S. K. (2020). Error-related potentials in passive brain-computer interfaces for quantitative assessment and adaptation of robotic systems during human-robot interaction (Doctoral dissertation, Technische Universität München).

Ehrlich, S., & Cheng, G. (2016). A neuro-based method for detecting context-dependent erroneous robot action. In 2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids) (pp. 477-482). IEEE.

Ferrez, P. & Millan, J. (2005) You are wrong!: automatic detection of interaction errors from brain waves. In Proceedings of the 19th International Joint Conference on Artificial Intelligence.

Fischer, K., Weigelin, H. M., & Bodenhagen, L. (2018). Increasing and adaptability. *Paladyn, Journal of Behavioral Robotics*, *9*: 95–109.

Flook, R., Anas Shrinah, Luc Wijnen, Kerstin Eder, Chris Melhuish, and Severin Lemaignan. (2019). On the impact of different types of errors on trust in human-robot interaction: Are laboratory-based HRI experiments trustworthy? Interaction Studies 20 (2019), 455–486

Fountaine, T., McCarthy, B., & Saleh, T. (2019). Building the AI-powered organization. Harvard Business Review

Glikson, E., & Woolley, A. W. 2020. Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14: 627–660.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology*, *2*(1), 141–165.

Goodrich, M. and A. C. Schultz, A. (2007). Human-robot interaction: a survey. Foundations and trends in human-computer interaction. 1 (3): 203-275

Hancock P.A., Billings D.R., Schaefer K.E., Chen J.Y.C., De Visser E.J., Parasuraman R. (2011).
A meta-analysis of factors affecting trust in human-robot interaction. Hum. Fact., 53 (5): 517-527

Hayes B, Scassellati B. (2015) Effective robot teammate behaviors for supporting sequential manipulation tasks. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE: 6374–6380.

Heaton, S., Teece, D., & Agronin, E. (2023). Dynamic capabilities and governance: An empirical investigation of financial performance of the higher education sector. *Strategic Management Journal, 44*(2), 520-548.

Heerink M., Kröse B., Evers V., Wielinga B. (2010). Assessing acceptance of assistive social agent technology by older adults: The almere model. International Journal of Social Robotics, 2, 361–375

Helfat, C. & Peteraf, M. (2009). Understanding dynamic capabilities: progress along a developmental path. *Strategic Organization*. 7(1): 91-102.

Helfat, C. & Peteraft, M. (2015). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. Strategic Management Journal. 36: 831-850.

Helfat CE, Winter SG (2011) Untangling dynamic and operational capabilities: Strategy for the (n)ever-changing world. *Strategic Management Journa*l. 32(11):1243–1250.

Helfat CE, Finkelstein S, Mitchell W, Peteraf MA, Singh H, Teece DJ, Winter SG, eds. (2007) Dynamic Capabilities: Understanding Strategic Change in Organizations (Blackwell Publishing, Oxford, UK).

Hopko, S., Wang, J., Mehta, R. (2022). Human factors considerations and metrics in shared space human-robot collaboration: A systematic review. Frontiers in Robotics and AI. 9

Ikemoto, S., Ben Amor, H., Minato, T., Jung, B., & Ishiguro, H. (2012). Physical human–robot interaction: Mutual learning and adaptation. IEEE Robotics Automation Magazine, 19(4), 24–35.

Kaplan, A. D., Sanders, T., & Hancock, P. A. (2019). The Relationship Between Extroversion and the Tendency to Anthropomorphize Robots: A Bayesian Analysis. *Frontiers in Robotics and AI*, *5*, 135.

Kim SK, Kirchner EA, Stefes A, Kirchner F. Intrinsic interactive reinforcement learning–Using error-related potentials for real world human-robot interaction. Scientific reports. 2017; 7(1):1–16.

Kirchner, E. A., Tabie, M. & Seeland, A. (2014). Multimodal movement prediction - towards an individual assistance of patients. *PLoS ONE.* **9**, e85060

Kump, B, Engelmann, A., Kessler, A., Schweiger, C. (2019). Toward a dynamic capabilities scale: measuring organizational sensing, seizing, and transforming capacities. Industrial Corporate Change. 28(5): 1149-1172

Letheren, K., Kuhn, K. A. L., Lings, I., & Pope, N. K. L. (2016). Individual difference factors related to anthropomorphic tendency. *European Journal of Marketing*, *50*(5/6), 973–1002

Leusmann, J., Wang, C., Gienger, M., Schmidt, A., Mayer, S. (2023). Understanding the uncertainty loop of human-robot interaction. arXivpreprint arXiv:2303.07889

Li, D., Liu, J. (2014). Dynamic capabilities, environmental dynamism, and competitive advantage: evidence from China. Journal of Business Research. 67(1): 2793-2799

Li, J. & Lee, Y. (2023). To disclose or not? Understanding employees' uncertainty and behavior regarding health disclosure in the workplace: A modified sociological approach. *International Journal of Business Communication*. 60(1): 173-201.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points". Annual Review of Psychology. 51: 405-444.

Luo, X., Qin, S., Fang, Z., & Qu, Z. (2021). Artificial intelligence coach for sales agents: Caveats and solutions. Journal of Marketing, 85(March), 14–32.

Mas, A., & Moretti, E. (2009). Peers at work. American Economic Review, 99(1), 112–145.

Memar, A. (2019). Enhancing human-robot interaction with passive brain-computer interfaces: A
neuroergonomic approach. Dissertation University at Buffalo

Meyer, G., Adomavicius, G., Johnson, P. E., Elidrisi, M., Rush, W. A., Sperl-Hillen, J. A. M., & O'Connor, P. J. (2014). A machine learning approach to improving dynamic decision making. Information Systems Research, 25(2), 239–263.

Miller, D., Minichilli, A., & Corbetta, G. (2013). Is family leadership always beneficial? Strategic Management Journal, 34(5): 553–571.

Mirnig, N., Gerald Stollnberger, Markus Miksch, Susanne Stadler, Manuel Giuliani, and
Manfred Tscheligi. 2017. To Err Is Robot: How Humans Assess and Act toward an Erroneous Social Robot. Frontiers in Robotics and AI 4 (2017).

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology, 48(2), 229-240.

Mori, M. (1970). The uncanny valley. *Energy* **7**, 33–35.

Mutlu B, Shiwa T, Kanda T, Ishiguro H, Hagita N. (2009) Footing in human-robot conversations:
how robots might shape participant roles using gaze cues. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. 61–68.

Natarajan, M., Gombolay, M. (2020). Effects of anthropomorphism and accountability on trust in human robot interaction Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction. 33-42

Natarajan, M., Seraj, E., Altundas, B., Paleja, R., Ye, S., Chen, L., Jensen, R., Chang, K., Gombolay, M. (2023). Human-robot teaming: Grand challenges. Current Robotics Reports. 4: 81-100.

Nikolaidis, S., Shah, J. (2013). Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategyProceedings of the 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI) 33-40

Nikolaidis, S., Hsu, D., and Srinivasa, S. (2017). Human-Robot Mutual Adaptation in Collaborative Tasks: Models and Experiments. *International Journal of Robotics Research*. 36, 618–634.

Peteraf, M., Di Stefano, G., Verona, G. (2013). The elephant in the room of dynamic capabilities: bringing two diverging conversations together. Strategic Management JOurnal. 34: 1389-1410.

Popper, K. R. 1957. ***T****he poverty of historicism*. Boston, MA: Beacon Press.

Pumplun, L., Tauchert, C., & Heidt, M. (2019). A new organizational chassis for artificial intelligence-exploring organizational readiness factors. In: Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden

Reagans, R. (2011). Close encounters: Analyzing how social similarity and propinquity contribute to strong network connections. Organization Science, 22(4), 835–849.

Rocha, M., da Silva, H., Morales, A., Sarkadi, A., Panisson, A. (2023). Applying theory of mind to multi-agent systems: A systematic review. *Intelligent Systems*. 367-381.

Rossi, A., Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters. 2017. How the Timing

and Magnitude of Robot Errors Influence Peoples' Trust of Robots in an Emergency Scenario. In International Conference on Social Robotics. Springer, 42–52.

Salem, M. Lakatos, G., Amirabdollahian, F., and Dautenhahn, K. 2015. Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. 141–148.

Salvato, C. & Vassal, R. (2018). The sources of dynamism in dynamic capabilities. Strategic Management Journal. 39: 1728-1752.

Santana, M., & Cobo, M. J. 2020. What is the future of work? A science mapping analysis. European Management Journal, 38: 846–862.

Schaefer, K. (2016). Measuring trust in human robot interactions: development of the "trust perception scale-HRI". Robust Intelligence and Trust in Autonomous Systems. 191-218.

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. Clinical neurophysiology, 118(1), 98-104.

Schoemaker, P., Heaton, S., & Teece, D. (2018). Innovation, dynamic capabilities, and leadership. California Management Review, 61, 15–42.

Sisbot, E., Marin-Urias, L., X. Broquere, X., D. Sidobre, D., R. Alami, R. (2010). Synthesizing robot motions adapted to human presence. *International Journal of Social Robotics*, 2(3): 329-343

Steinbauer.S. 2013. A Survey about Faults of Robots Used in RoboCup. In RoboCup 2012: Robot Soccer World Cup XVI. 344–355.

Stiber, M. & Huang, C. (2020). Not all errors are created equal: exploring human responses to robot errors with varying severity. In Companion Publication of the 2020 International

Conference on Multimodal Interaction. 97–101.
Sun, C., Shi, Z., Liu, X., Ghose, A., Li, X., & Xiong, F. (2019). The effect of voice AI on consumer purchase and search behavior. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3480877
Sung, J.-Y., R. E. Grinter, and H. I. Christensen. (2010). Domestic robot ecology. International Journal of Social Robotics 2 (4):417–29
Szafir D, Mutlu B. (2012). Pay attention! Designing adaptive agents that monitor and improve user engagement. In: Proceedings of the SIGCHI conference on human factors in computing systems. 11–20.
Taylor, F.W.. 1911. The principles of scientific management. New York: Harper & Bros.
Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28, 1319–1350.
Teece DJ, Pisano G, Shuen A (1997) Dynamic capabilities and strategic management. *Strategic Management Journal*. 18(7):509–533.
Teece, D. (2016). Dynamic capabilities and entrepreneurial management in large organizations: Toward a theory of the (entrepreneurial) firm. 86: 202-216. *European Economic Review*.
Thompson, J.D. 1967. Organizations in action. New York: McGraw-Hill.
Tong, S., Jia, N., Luo, X., Fang, Z. (2021). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. Strategic Management Journal.
Tong, Y., Liu, H., and Zhang, Z. (2024). Advancements in humanoid robots: A comprehensive review and future prospects. IEEE/CAA J. Autom. Sinica, 11 (2): 301–328
You, A., Kim,J., Lee, S., Kamat, V., and Robert, L. (2018). Enhancing perceived safety in human–robot collaborative construction using immersive virtual environments, *Autom. Construct.*, 96: 161–170.
Venkatasubramanian, S., Bliss, N., Nissenbaum, H., & Moses, M. (2020). Interdisciplinary approaches to understanding artificial intelligence's impact on society. *arXiv preprint arXiv:2012.06057*.
Whetten, D. A. 1989. What constitutes a theoretical contribution? Academy of Management Review, 14(4): 490-495.
Wiese E, Metta G, Wykowska A. (2017). Robots as intentional agents: using neuroscientific methods to make robots appear more social. Frontiers in psychology. 8:1663.
Wollersheim, J. & Heimeriks, K. (2016). Dynamic capabilities and their characteristic qualities: Insights from a lab experiment. Organization Science. 27(2): 233-248.
Zacharaki, A., Kostavelis, I., Gasteratos, A., and Dokas, I. (2020). Safety bounds in human robot interaction: A survey," *Safety science*, vol. 127, p. 104667.
Zhang, X. & You, X. (2020). The impact of perceived risk on consumers' cross-platform buying behavior. *Frontiers in Psychology*. 11: 592246
Zollo M, Winter SG (2002) Deliberate learning and the evolution of dynamic capabilities. Organization Science. 13(3):339–351.