Dec 11th, 12:00 AM

Dynamics and Impacts of Human-Algorithm Consensus in Logistics Scheduling: Evidence from A Field Experiment

Lingli Wang
Beijing University of Posts and Telecommunications, wang.ll@bupt.edu.cn

Ding Wu
School of Information Technology and Management, 03129@uibe.edu.cn

Nina Huang
University of Miami, nihuang@uh.edu

Hongshuyu Deng
JD.com.Inc, denghongshuyu@jd.com

Xiaotian Zhuang
JD.com.Inc, zhuangxiaotian@jd.com

Dynamics and Impacts of Human-Algorithm Consensus in Logistics Scheduling: Evidence from A Field Experiment

Short Paper

Lingli Wang
Beijing University of Posts and Telecommunications
Beijing, China, 100876
wang.ll@bupt.edu.cn

Ding Wu
University of International Business and Economics
Beijing, China, 100029
03129@uibe.edu.cn

Ni Huang
University of Miami
Coral Gables, FL 33146
nhuang@miami.edu

Hongshuyu Deng
JD.com.Inc
Beijing, China, 100176
denghongshuyu@jd.com

Xiaotian Zhuang
JD.com.Inc
Beijing, China, 100176
zhuangxiaotian@jd.com

Abstract

Algorithms are being implemented to aid human decision-making and most studies on human-algorithm interactions focus on how to improve human-algorithm cooperation. However, excessive reliance on algorithms in decision-making may hinder the complementary value of humans and algorithms. There is a lack of empirical evidence on the impacts of human-algorithm consensus in collaborative decision-making. To address this gap, this paper reports a large-scale field experiment conducted by one of China's largest logistics firms in the context of route scheduling. The experiment involved assigning routes to either a treatment group, where algorithms and human operators collaborated in decision-making, or a control group, where human operators made decisions independently. We plan to collect data to evaluate the effects of algorithm implementation and to analyze the patterns and effects of human-algorithm consensus in a long-term cooperation. Our study aims to contribute to the literature on human-algorithm interactions in operational decisions.

Keywords: Operational decision, human-algorithm cooperation, field experiment

Introduction

In recent years, the use of artificial intelligence (AI) algorithms has become increasingly popular for revolutionizing or facilitating business operations (Sun et al., 2022). With the ability to process large amounts of historical data, these algorithms have achieved exceptional performance in various tasks, such as assigning pick tasks (Bai et al., 2021), providing packing instructions (Sun et al., 2022), recommending quality improvement actions (Senoner et al., 2022), and managing production scheduling tasks (Chien et al., 2020). Online experiments have also demonstrated that integrating algorithms into human decision-making processes can improve the quality of final judgments or performance (Fugener et al., 2021).

Despite the advantages of AI algorithms, evidence from both real business settings and lab experiments reveals the existence of algorithm aversion phenomenon when humans interact or cooperate with algorithms. For example, Sun et al. (2022) found information deviation and complexity deviation when workers executed algorithm-generated order packing instructions. Dietvorst et al. (2015) demonstrated that individuals are reluctant to accept advice from algorithms, especially after they have seen algorithms err. Although algorithms may not work perfectly, they still have advantages over humans, particularly in making decisions that require a comprehensive understanding of data (Tong et al., 2021). Thus, scholars suggest that there are systematic complementarities between humans and advanced algorithms, with algorithms being able to handle and learn from a large amount of decision-supporting data, while humans contribute new knowledge to decision-making (Fugener et al., 2021). Given the potential benefits of human-algorithm complementarity and the existence of algorithm aversion, there is significant academic interest in exploring how to optimize designs to improve human-AI cooperation (Dietvorst et al., 2018; Castelo et al., 2019).

However, the literature on technology adoption suggests that users are more likely to adopt or rely on new technologies after having positive interactions with them (Komiak and Benbasat, 2006; Esmaeilzadeh, 2019). As users become more familiar with algorithms, they are likely to update their attitudes towards them and rely on them more (Cao and Zhang, 2021). According to Fugener et al. (2022), the assistance of algorithms may dominate individuals' final decision, resulting in a decrease in the contribution of unique human knowledge. Therefore, in the context of long-term human-algorithm collaboration, humans' reliance on algorithms may continue to increase, potentially hindering their contribution of new knowledge or external information in decision-making, leading to less effective collaboration. To the best of our knowledge, there is a lack of empirical evidence on the patterns and impacts of human-algorithm consensus (i.e., the extent to which humans rely on algorithms to make decisions) in a long-term collaboration context. Thus, this paper aims to address the following research questions:

How does the implementation of algorithms affect the final decision performance in a long-term collaboration*?*

How does the human-algorithm consensus dynamically change over time in a long-term collaboration?

What is the relationship between the human-algorithm consensus and the final decision performance?

In order to address the research questions, we report a large-scale field experiment conducted by one of the largest logistics firms in China, in the context where humans alone or in collaboration with algorithms to make logistics scheduling decisions. Logistics scheduling is a crucial and intricate decision-making process in operations management, which involves allocating available resources to transport products (i.e., items to be delivered to customers) while minimizing costs and meeting constraints of delivery time, product characteristics, and vehicle availability (Chen and Vairaktarakis, 2005; Liu et al., 2020). In such a decision-making context, algorithms and humans possess unique strengths. Algorithms have the ability to analyze vast amounts of historical data and learn from various cases to identify relatively optimal solutions that satisfy complex constraints (Tong et al., 2021). On the other hand, human operators often make decisions heuristically based on their prior decision-making experience. Additionally, human operators can actively observe external information overlooked by algorithms and provide flexible responses to unique cases (Sun et al., 2022).

In the month-long field experiment, two sets of comparable scheduling routes were assigned to either the treatment or control group. For scheduling routes in the control group, human operators manually assign available vehicles to transport certain products. In the treatment group, the partner firm implemented an advanced scheduling algorithm[1] to first generate recommendations on how to assign available vehicles to specific products with the lowest cost, and human operators then decide whether to accept or modify

[1] The partner firm used the LightGBM (Light Gradient Boosting Machine) algorithm to generate transportation plans. Please see more introduction about the algorithm at https://lightgbm.readthedocs.io/en/v3.3.2/

parts or all of the algorithms' advice. To test our hypotheses, we plan to collect data about route-level, human operator-related, and human-algorithm interaction-related information from the partner firm. We will first conduct randomization checks to ensure the execution of the field experiment was successful. We will then conduct regressions to confirm the value of algorithm implementation in improving final scheduling decision performance. Next, based on detailed data on human-algorithm interactions, we will examine the dynamic changes in human-algorithm consensus and analyse how it affects the final decision performance in the long-term collaboration.

Our study aims to contribute to the literature on human-algorithm interactions in several ways. First, we plan to evaluate the economic value of advanced algorithms in improving the performance of logistics scheduling decisions using field evidence from a large-scale experiment with a logistics firm in China. Second, and more importantly, we seek to capture how the pattern of human-algorithm consensus changes dynamically in a long-term human-algorithm cooperation context. Third, we will examine the relationship between human-algorithm consensus and the final decision-making performance, providing evidence for the negative effects of humans' over-reliance on algorithms. Our study has the potential to provide practical implications for designing interventions that optimize the cooperation between humans and algorithms.

Literature Review

Algorithms in Operations Management

Our research is situated within the wider literature on algorithms in operations management. With the increasing availability of data and the development of analytical technologies, algorithms are being widely implemented to facilitate or support operational decision-making. Previous studies have demonstrated the value of certain algorithms, such as those implemented to automatically assign pick tasks in warehouses (Bai et al., 2021), provide packing instructions for workers (Sun et al., 2022), recommend suitable quality improvement actions in manufacturing (Senoner et al., 2022), and handle production scheduling tasks (Chien et al., 2020).

Specifically, in operations management, logistics scheduling is a complicated operational problem that involves planning how to effectively deliver products or services through logistics workflows or pipelines while considering time constraints, vehicle constraints, cost constraints, and customer requirements. As a complicated optimizing problem, previous studies on logistics scheduling mainly focus on proposing designs to improve algorithm performance in specific contexts (Bramel and Simchi-Levi, 1997; Chen and Vairaktarakis, 2005; Chen and Lee, 2008; Liu et al., 2020; Wu et al., 2023). However, as pointed out by Sun et al. (2022) and Donselaar et al. (2010), algorithms may sometimes ignore important contextual constraints, leading to suboptimal decisions. Meanwhile, humans can actively collect extra-contextual information, which can support improving the quality of the decision-making process (Sun et al., 2022). Thus, in practice, platforms often rely on humans and algorithms together to make final logistics scheduling decisions. Our study contributes to this stream of literature by providing empirical evidence on how humans and algorithms work together to make final decisions, particularly by revealing the dynamic pattern of human-algorithm collaboration, which we define as human-algorithm consensus in this study.

Human-algorithm Interactions

Our study also aims to contribute to the literature on human-algorithm interactions. Recent literature has demonstrated the phenomenon of algorithmic aversion in various contexts. For example, studies show that customers are less likely to buy financial products promoted by AI chatbots (Luo et al., 2019) and workers tend to deviate from algorithm suggestions when making decisions (Sun et al., 2022). Disclosing the use of AI to provide feedback to employees results in negative perceptions among employees (Tong et al., 2021), and even when individuals know that algorithms outperform humans, they may still be reluctant to follow algorithmic advice if they know that the algorithm make errors (Dietvorst et al., 2015). Meanwhile, scholars suggest that human-algorithm collaborations can add value in many contexts by taking advantage of the strengths of both algorithms and humans (Luo et al., 2021; Sun et al., 2022). For example, the AI-human coach assemblage outperforms either the AI or human coach alone when performing sales agent training tasks (Luo et al., 2021). Moreover, when humans and algorithms work together to perform classification tasks, they can outperform the algorithm when the algorithm delegates some work to humans (Fugener et al., 2022).

Therefore, scholars are increasingly interested in exploring effective designs to change individuals' perceptions of algorithms and enhance human-algorithm cooperation. For example, some studies suggest giving people some control over an algorithm (Dietvorst et al., 2018) or increasing a task's perceived objectivity and algorithms' perceived affective human-likeness (Castelo et al., 2019) are effective designs to mitigate algorithm aversion. While almost all studies assume the positive effects of human-AI cooperation, Fugener et al. (2021) caution that algorithms may facilitate humans' decision-making performance at the individual level, but also leads to a convergence towards similar responses and ultimately decreases the contribution of unique human knowledge at the group level. Our study aims to contribute to this literature by leveraging a large-scale field experiment to explore how the extent of human-algorithm consensus affects the final decision performance in long-term human-algorithm cooperation. We are particularly interested in investigating the potential negative effects of humans' over-reliance on algorithms in decision-making.

Hypotheses Development

In dealing with complicated decision-making tasks (e.g., logistics scheduling), with the rapid proliferation of available data and the rise in the number of constrains, it is increasingly difficult for humans to manually make such decisions effectively (Tong et al., 2021). Moreover, as noted by Lu and Yin (2021), individuals may not always recognize their limitations in decision-making, and they tend to rely on heuristics and past experience rather than analyzing available data to make decisions. This can result in under utilization of available information for making better decisions. Conversely, algorithms have the potential to quickly and accurately analyze large amounts of data while satisfying complex constraints due to advancements in computing power (Tarafdar et al., 2019), leading to better predictions and decisions than those made by humans (Dietvorst et al., 2015; Tong et al., 2021). Building on the existing research that examines the value of algorithm implementations in various settings, we propose the following hypothesis.

H1: Implementation of algorithms in logistic scheduling decisions improves decision performance compared to that of decisions made by humans alone.

Although algorithms can provide complementary value to human decision-making process, recent literature has demonstrated the phenomenon of algorithm aversion, particularly when individuals have limited knowledge or are unfamiliar with algorithms (Castelo et al., 2019; Luo et al., 2019). However, the literature on technology adoption suggests that users gradually gain familiarity with new technologies, systems, or designs through interactions with them (Gefen et al., 2003; Komiak and Benbasat, 2006). When such new technologies or systems perform well, familiarity increases users' trust in them (Komiak and Benbasat, 2006; Esmaeilzadeh, 2019). In the context of AI adoption, via a field experiment, Cao and Zhang (2021) found that forcing workers to accept the assistance of algorithms is helpful in inducing them to update their beliefs about algorithms, which in turn encourages them to form the habit of using algorithms. Their experiment results also suggest that, even without any external intervention, workers themselves become more familiar with the algorithm and are more likely to adopt and use it (Cao and Zhang, 2021). Therefore, we propose that

H2: The extent of human-algorithm consensus has a positive relationship with the number of human-algorithm interactions.

Algorithms are capable of handling complex data analysis tasks, while humans are able to capture important contextual information overlooked by algorithms (Tong et al., 2021). As humans and algorithms possess complementary skills in decision-making (Fugener et al., 2021), human-algorithm cooperation can lead to improved decision performance. However, Fugener et al. (2021) found that the implementation of algorithms can decrease users' unique knowledge contribution, especially when workers become too reliant on algorithms to make complex decisions. When this occurs, there is a high level of consensus between humans and algorithms, which undermines their complementary value, particularly for dealing with certain decision cases that need human observed external information (e.g., encountering extreme weather and having certain types of products to deliver). Consequently, excessively high levels of human-algorithm consensus may lead to lower decision performance. Therefore, we propose the following hypothesis:

H3: There is a inverted U-shape relationship between human-algorithms consensus and the performance of scheduling decisions.

Research Context and Experimental Design

To test our hypotheses and explore potential underlying mechanisms, we plan to analyse the data from a large-scale field experiment conducted by one of the largest logistics firms in China in December 2022. The partner firm has a logistic scheduling system that collects and organizes customer-, order-, scheduling route-, and vehicle-related information. The firm manages its logistics scheduling decisions by route (i.e., making decisions of how to assign available vehicles to transport all the goods from one place to another with the lowest cost while satisfying constraints of delivery time, product characteristics, available vehicles, etc.) and scheduling routes that share the same node (a start or end place) form a route network. Each route typically involves several decision tasks (i.e., selecting a suitable vehicle for a specific product) and multiple products can be assigned to the same vehicle.

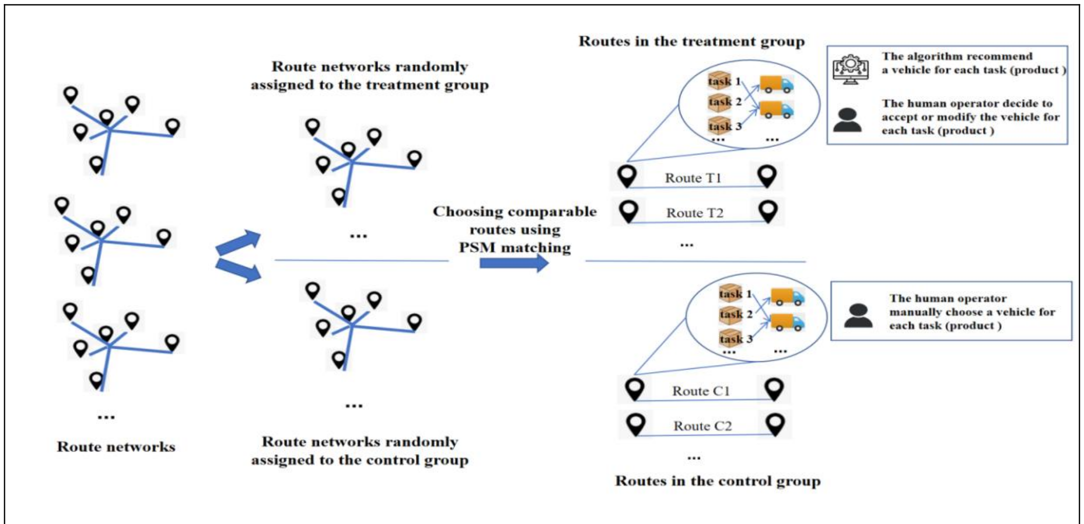

Figure 1. Experimental Design

Before the experiment, human operators relied on their experience to select the most suitable vehicle for all products. In the experiment, separate route networks were randomly assigned to either the treatment or control group. That is, the partner firm conducted experiment randomization at the route network level. Then, in order to ensure the comparability of routes in the treatment and control groups, the firm used propensity score matching (PSM) to select two sets of comparable routes from the treatment and control groups, based on important route-level information such as distance, number of available vehicles, number of historic tasks, etc. For all observation routes in the treatment group, an advanced logistics scheduling algorithm (i.e., the LightGBM algorithm) was implemented to generate a transportation plan for distributing specific products to the most suitable available vehicle. Human operators then checked the plan and had the option to accept or modify it by task. If the algorithm recommended a suitable vehicle for specific products, the human operators accepted the plan. However, if the algorithm failed to recommend a suitable vehicle, the human operators modified the plan accordingly. For observation routes in the control group, the algorithm did not recommend transportation plans, and human operators made decisions manually. The experiment lasted for about one month, covering more than 3,000 scheduling routes in more than one hundred cities in China. Figure 1 illustrates the detail of the field experimental design.

Data and Analysis Plan

We plan to extract data from the partner firm's database, which is expected to contain route-level basic information, human operators' basic information, and details about the operator-algorithm interactions during and after the field experiment. Table 1 summarizes the descriptions of variables we plan to construct.

Variable	Description
Treatment	For observational routes in the treatment group, *Treatment* = 1; for observational routes in the control group, *Treatment* = 0.
Num vehicle	The number of available vehicles for a route.
Historical task	The number of historical tasks for a route in the month prior to the start of the experiment.
Historical task variation	The standard deviation of the number of historical tasks a route had in one month before the experiment.
Recommended transportation plan	Details of the transportation plans the algorithm recommended for human operators. For observational routes in the control group, this variable will not be available since no recommendations were made by the algorithm.
Final transportation plan	The final transportation plan for the treatment group includes both the algorithm's recommendations and any modifications made by human operators, whereas in the control group, the final transportation plan only includes the decisions made by human operators without any algorithmic input.
Algorithm human consensus	The proportion of recommended transportation plans that were accepted by human operators without modification, calculated based on the differences between the *recommended transportation plan* and the *final transportation plan*.
The actual cost	The cost calculated based on the final submitted transportation plans co-contributed by the algorithm and human operators.
Algorithm transportation plan cost	The cost calculated based on the algorithm generated transportation plans.
The minimum cost	The minimum possible cost calculated for each route at the post-evaluation stage after all products have been delivered to their respective destinations.

Table 1. Descriptions of Variables in the Study

Prior to our primary analysis, we will perform randomization checks to ensure the comparability of observational routes in the treatment and control groups. Subsequently, we will develop an ordinary least square regression model, as specified in Equation (1) to capture the direct effects of the implementation of the scheduling algorithm.

$$Outcome\ Variable_{ijt} = \alpha_0 + \alpha_1 Treatment_i + HControls_{it} + JD_t + KR_i + LO_j + \varepsilon_{ijt} \quad (1)$$

Where i denotes scheduling routes, j denotes operators, and t denotes observation time. Outco*me Variable*$_{ijt}$ represents *the actual cost* for each route and the difference between actual and minimum costs for each route in the treatment or control group. The coefficient of $Treatment_i$ will indicate whether the implementation of the scheduling algorithm has a significant effect on the final decision performance. To account for potential confounding factors, we will include a set of control variables that capture time-relevant route-level information, the effect of each day during our experiment (i.e., D_t), route-fixed effects (R_i), and operator-fixed effects (O_j).

Further, as specified in Equations (2) and (3), we will explore how *human-algorithm consensus changes along with the increase in the number of human algorithm interactions, and* the relationship between the *human-algorithm consensus* and the final decision performance using observations in the treatment group.

$$Human\ algorithm\ Consensus_{ijt} = \alpha_0 + \alpha_1 Human\ algorithm\ interaction_j + HControls_{it} + JD_t + KR_i + LO_j + \varepsilon_{ijt} \quad (2)$$

$$Outcome\ Variable_{ijt} = \alpha_0 + \alpha_1 Human\ algorithm\ consensus^2_{ijt} + \alpha_2 Human\ algorithm\ consensus_{ijt} + HControls_{it} + JD_t + KR_i + LO_j + \varepsilon_{ijt} \quad (3)$$

Where i denotes scheduling routes, j denotes operators, and t denotes observation time. Outcome Variableijt represents the actual cost for each route and the difference between actual and minimum costs for each route in the treatment or control group. *Human algorithm interaction* measures the number of times a human operator relying on the algorithm to make decisions. The coefficients of *Human algorithm consensus$_{ijt}$ and Human algorithm consensus$_{ijt}^2$* will suggest the relationship between the extent of human operators' reliance on the algorithm and the performance of the final decisions.

In addition, after the main analyses, we will examine the heterogeneous of the effects of the algorithm implementation and human-algorithm consensus by taking into account the moderating effects of the decision complexity for each route and the characteristics of human operators.

Discussion, Expected Contribution, and Limitation

Algorithms are increasingly being implemented to facilitate human decision-making in operations management. While algorithms have advantages in processing and understanding large amounts of data quickly and accurately, humans have the ability to capture information that algorithms may overlook and give flexible responses to specific contextual changes. Therefore, scholars are paying increasing attention to exploring designs that promote collaboration between humans and algorithms and mitigate the phenomenon of algorithm aversion. However, the implementation of algorithms may backfire if humans begin to over-rely on algorithms to make decisions. This study leverages data from a large-scale field experiment conducted by one of the largest logistics firms in China. The study aims to quantify the effects of algorithm implementation in a logistics scheduling context and, more importantly, to examine the dynamic changes in human-algorithm consensus and their impacts on final decision performance.

Through this study, we expect to make the following contributions. First, our study aims to extend the literature on human-algorithm interactions by evaluating the economic value of implementing advanced algorithms in a context (i.e., logistics scheduling) where humans and algorithms cooperate to make decisions using field evidence from a large-scale experiment. More importantly, we plan to explore how the extent of human-algorithm cooperation (i.e., human-algorithm consensus) changes dynamically along with the increase in the number of human-algorithm interactions. Taking a step further, we will examine the relationship between human-algorithm consensus and the final decision-making performance, revealing that humans' over-reliance on algorithms will bring about a reduction in the complementary value of human-algorithm cooperation.

Our study also has limitations. For example, we plan to analyze data from a one-month field experiment. It could not address possible periodic seasonal effects in the current study design. We mainly consider analyzing objective data extracted from the partner firm and we could not get information about human operators' subjective perceptions of and attitudes to the algorithm.

Acknowledgements

Lingli Wang acknowledges support from the Beijing University of Posts and Telecommunications Basic Scientific Research Program (2022RC21, 2023KJBF5) and National Natural Science Foundation of China (72201038).

References

Bai, B., Dai, H., Zhang, D. J., Zhang, F., and Hu, H. 2022. “The Impacts of Algorithmic Work Assignment on Fairness Perceptions and Productivity: Evidence from Field Experiments,” *Manufacturing & Service Operations Management* (24:6), pp. 3060-3078.

Bramel, J., and D. Simchi-Levi. 1997. *Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management*, Springer Verlag, New York.

Castelo, N., Bos, M. W., and Lehmann, D. R. 2019. “Task-dependent Algorithm Aversion,” *Journal of Marketing Research* (56:5), pp. 809-825.

Cao, X., and Zhang, D. 2020. “The Impact of Forced Intervention on AI Adoption,” Available at SSRN 3640862.

Chen, B., and Lee, C. Y. 2008. “Logistics Scheduling with Batching and Transportation,” *European Journal of Operational Research* (189:3), pp. 871-876.

Chen, Z. L., and Vairaktarakis, G. L. 2005. “Integrated Scheduling of Production and Distribution Operations,” *Management Science* (51:4), pp. 614-628.

Chien, C. F., Dauzere-Peres, S., Huh, W. T., Jang, Y. J., and Morrison, J. R. 2020. “Artificial Intelligence in Manufacturing and Logistics Systems: Algorithms, Applications, and Case Studies,” *International Journal of Production Research* (58:9), pp. 2730-2731.

Dietvorst, B. J., Simmons, J. P., and Massey, C. 2015. “Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err,” *Journal of Experimental Psychology: General* (144:1), pp. 114–126.

Dietvorst, B. J., Simmons, J. P., and Massey, C. 2018. “Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms if They Can (Even Slightly) Modify Them,” *Management Science* (64:3), pp. 1155–1170.

Esmaeilzadeh, P. 2019. “The Process of Building Patient Trust in Health Information Exchange (HIE): The Impacts of Perceived Benefits, Perceived Transparency of Privacy Policy, and Familiarity,” *Communications of the Association for Information Systems* (45), pp. 364-396.

Fugener, 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* (45:3), pp. 1527-1556.

Fugener, 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.

Gefen, D., Karahanna, E., and Straub, D. W. 2003. “Trust and TAM in Online Shopping: An Integrated Model,” *MIS Quarterly* (27:1), pp. 51-90.

Komiak, S. Y., and Benbasat, I. 2006. “The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents,” *MIS Quarterly* (30:4), pp. 941-960.

Luo, X., Qin, S., Fang, Z., and Qu, Z. 2021. “Artificial Intelligence Coach for Sales Agents: Caveats and Solutions,” *Journal of Marketing* (85:March), pp. 14–32.

Luo, X., Tong, S., Fang, Z., and Qu, Z. 2019. “Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases,” *Marketing Science* (38:6), pp. 937–947.

Senoner, J., Netland, T., and Feuerriegel, S. 2022. “Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing,” *Management Science* (68:8), pp. 5704-5723.

Sun, J., Zhang, D. J., Hu, H., Van Mieghem, J. A. 2022. “Predicting Human Discretion to Adjust Algorithmic Prescription: A Large-Scale Field Experiment in Warehouse Operations,” *Management Science* (68:2), pp. 846-865.

Liu, C., Feng, Y., Lin, D., Wu, L., and Guo, M. 2020. “Iot Based Laundry Services: An Application of Big Data Analytics, Intelligent Logistics Management, and Machine Learning Techniques,” *International Journal of Production Research* (58:17), pp. 5113-5131.

Lu, Z., and Yin, M. 2021. “Human Reliance on Machine Learning Models When Performance Feedback is Limited: Heuristics and Risks,” *In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1-16.

Tarafdar, M., Beath, C. M., and Ross, J. W. 2019. “Using AI to Enhance Business Operations,” *MIT Sloan Management Review* (60:4), pp. 37–44.

Tong, S., Jia, N., Luo, X., and Fang, Z. 2021. "The Janus Face of Artificial Intelligence Feedback: Deployment versus Disclosure Effects on Employee Performance," *Strategic Management Journal* (42:9), pp. 1600-1631.

Van Donselaar, K. H., Gaur, V., Van Woensel, T., Broekmeulen, R. A., and Fransoo, J. C. 2010. "Ordering Behavior in Retail Stores and Implications for Automated Replenishment," *Management Science* (56:5), pp. 766-784.