

In control or under control? Human–robot collaboration in warehouse order picking

Alexandros Pasparakis

Erasmus University Rotterdam, Rotterdam, The Netherlands

Jelle De Vries

*Department of Information Systems and Operations Management,
Tilburg University, Tilburg, The Netherlands, and*

René De Koster

Erasmus University Rotterdam, Rotterdam, The Netherlands

89

Received 19 March 2025
Revised 12 September 2025
11 December 2025
Accepted 2 February 2026

Abstract

Purpose – Warehouse operations are rapidly becoming more robotized to increase performance and reduce costs. A key innovation in the field is human–robot collaboration (HRC). Despite the essential role of human behavior in operational performance, human factors in HRC are still understudied in Operations Management literature. To address this gap, we conducted a unique real-effort experiment in a warehouse especially erected for this study and ground our analysis in Self-Determination Theory, Leader-Member Exchange Theory and Regulatory Focus Theory.

Design/methodology/approach – The experiment compares the objective outcomes of collaborative productivity, collaborative accuracy and human pick time between a configuration with the human leading the robot and a configuration with the human following the robot. Additionally, we investigate the behavioral mechanism governing human reactions to the novel collaboration with robots.

Findings – We find that human leading allows for superior collaborative order picking productivity, while human following allows for greater collaborative order picking accuracy. Especially when picking involves traveling between locations, human leading results in shorter pick times. We additionally establish a prevention focus as the factor that allows workers to bridge the productivity gap between the two setups.

Research limitations/implications – A limitation of the lab experiment is its short duration, which necessitates follow-up research in field experiments.

Originality/value – These insights will help warehousing service providers to tailor collaborative robotic solutions based on quantifiable trade-offs between system speed and accuracy.

Keywords Human factors, Behavioral operations management, Robotic warehouse systems, Human–robot collaboration

Paper type Research article

1. Introduction

Robots have been used in industry for decades. The latest generation of fully autonomous robots operates independently, enabling them to explore, navigate and manipulate their environment without human supervision (Ben-Ari and Mondada, 2017; Nayyar and Kumar, 2020; Azadeh *et al.*, 2019, 2025). Warehouses have served as leading incubators for the design and application of autonomous robots, as their structured environment makes them well-suited for robotic material handling (Azadeh *et al.*, 2019; Fragapane *et al.*, 2021). As a result, warehousing service providers have increasingly adopted autonomous robots, following advances in safe navigation close to humans (Ben-Ari and Mondada, 2017).



© Alexandros Pasparakis, Jelle De Vries and René De Koster. Published in *Logistics Research*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at [Link to the terms of the CC BY 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Logistics Research
Vol. 19 No. 1, 2026
pp. 89–113
Emerald Publishing Limited
e-ISSN: 1865-0368
p-ISSN: 1865-035X

DOI 10.1108/LORE-03-2025-0028

However, robots are not replacing human workers entirely. These “automatically controlled, reprogrammable, multipurpose manipulators” (International Organization for Standardization, 2012) were integral enablers for automation in the third industrial revolution. In Industry 4.0, they attained the pivotal ability to sense the environment, learn and adapt in real time, allowing them to become truly autonomous (Nayyar and Kumar, 2020; Neumann *et al.*, 2021). Nevertheless, when tasks cannot be fully automated (like grabbing product units from a location), human involvement is still required. Rather than using machines as tools, humans increasingly share the work floor and operate in collaboration with autonomous robots. This shift has led to the introduction of collaborative robots or “cobots”, designed to complement human labor rather than replacing it (Matheson *et al.*, 2019; Lorson *et al.*, 2022).

Despite advances in computer science and engineering, Operations Management research on robotics implementations is not abundant. As the dawn of Industry 5.0 moves industry further towards human–machine integration, the importance of studying the human factor for effectively designing, managing and deploying robots alongside human workers is crucial (Donohue *et al.*, 2020; Olsen and Tomlin, 2020; Sgarbossa *et al.*, 2020). Human–robot collaborative performance is dictated not only by individual human or robot performance but also by their shared collaborative performance. However, unlike robots, human performance is state-dependent, meaning that humans may adjust their work rates in response to the performance of their coworkers (e.g. in serial production lines: Cantor and Jin, 2019; Doerr *et al.*, 2002; Powell and Schultz, 2004; Schultz *et al.*, 1998). These dynamics are particularly relevant to warehouse order picking, where cobots assist human workers by handling repetitive tasks (e.g. transport) while humans focus on more complex tasks (e.g. order picking) (Azadeh *et al.*, 2019; Fragapane *et al.*, 2021; Gutelius and Theodore, 2019; Lorson *et al.*, 2022).

This study examines two behavioral factors that may shape the outcomes of human–robot collaboration (HRC). In human–robot collaborative order picking, productivity depends on both the joint actions of humans and robots and the tasks performed separately. During collaboration, humans and robots work together, while the robot later operates independently to deliver completed orders. Because robots handle transport, human pickers reduce travel, potentially making cobotic systems more productive than traditional order picking. A first behavioral implication is caused by the fact that humans and robots often alternate between the roles of “leading” or “following” one another (Wang *et al.*, 2017, 2020; Yanco and Drury, 2004). Depending on the specific setup, the robot may hinder or assist the human, resulting in varying collaborative performance outcomes that remain unexplored. Second, human behavior may be affected by the psychological construct known as prevention focus, which encapsulates the level of vigilance that an individual exerts to handle their responsibilities (Higgins, 1998). As such, we propose that when workers perceive robots as a hindrance, a higher prevention focus may compensate for any performance limitations inflicted by the robot.

We investigate these operational and behavioral implications of HRC in the context of warehouse order picking, as it provides a natural setting for the leading/following relationship to manifest. Order picking cobotic setups in the market include instances of humans leading (e.g. Toyota-BT, 2025), following/supporting the robot (e.g. 6 River Systems, 2019) or a hybrid combination where the roles alternate (e.g. Locus Robotics, 2025). Full automation of the order picking process remains impractical due to the lack of standardization, the variety of products, storage locations, order composition and requirements in terms of stacking, order handling and speed. Market projections suggest rapid growth in warehouse collaborative robotics (Statista, 2024). Understanding what constitutes effective HRC in warehouse order picking is both relevant and timely for future operational strategy.

With our study, we therefore address two key research questions: (1) How do *human leading vs human following* HRC dynamics impact order picking performance in terms of collaborative productivity, collaborative accuracy and human pick time? (2) What behavioral factor influences human responses to HRC? As such, we contribute to the literature of logistics research and operations management in three ways. First, we introduce the distinction of *human leading vs human following* relationship dynamics in HRC, motivated by the

manufacturing and robotics design literature on active/passive cobotic setups. Second, we propose a prevention focus as the key factor influencing the state-dependent human behavior in collaborative robotics settings. Third, we run an elaborate physical, real-effort experiment (i.e. where participants must perform actual, measurable tasks, rather than simply making choices, [Charness and Kuhn, 2011](#)), which is a novel approach compared to the simulated HRC setups commonly employed by studies in this domain. The findings provide insights into assessing new collaborative robotic solutions, enabling warehousing service providers to optimize human workforce for maximal operational efficiency.

2. HRC dynamics

HRC dynamics are shaped by the assignment of active (leading) and passive (following) roles to the collaboration partners. Across the stream of literature on robot design for human–robot interaction (HRI) in production lines, task-sharing robots can be characterized by their autonomy level ([Bruemmer et al., 2005](#); [Yanco and Drury, 2004](#)). A robot’s autonomy can be taxonomized by the level of human assistance they require to complete tasks, ranging from performing tasks almost independently (active robot) to relying heavily on human assistance (non-active/passive robot). Nowadays, the sub-stream of HRC has emerged. Under a collaborative relationship, the robot can now be the intervening agent and humans can be the “active” or “non-active”/“passive” agent according to their involvement in the task. Therefore, the concept of “agent autonomy” in HRC describes the extent to which the robot’s actions are directly determined by the human agent or vice versa. The relationship dynamics are closely related to “leader-follower” relationships and each partner’s behavior can range from inactive (resting), supportive (following), to active (leading) ([Wang et al., 2017, 2020](#)). As the inactive state has no collaboration implications, we focus on the opposing states of leading vs following.

In car manufacturing, an example of a supportive/following cobot assists active/leading human workers by handing them tools and parts they require during the assembly process. In the same assembly process, an active/leading cobot may lift, place or manipulate heavy loads, then hold them and wait for the supportive/following human to mount them together. Similarly, warehouse order picking provides valuable context to examine the leader-follower HRC dynamics. In low-level picker-to-parts order picking (i.e. retrieving products from low-level storage positions to fulfil specific customer orders), specific types of cobots collaborate with human pickers, adopting either a leading or a supportive role, causing the pickers to assume the complementary supportive or leading role, respectively. While some configurations allow for hybrid/adaptive collaboration (the two partners may alternate leading/following roles), we focus on distilling the opposing configurations where each partner takes full responsibility for a role.

The difference between the two setups lies in the way that movement from one pick location to the next is initiated. In a leading/active role, the cobot autonomously navigates to pick locations indicated by tasks retrieved from the warehouse management system. Subsequently, the picker interacts with the cobot at the pick location (reached either by visiting or by following the robot), puts the requested item in the appropriate load carrier and completes the pick, before following the cobot to the next pick location or searching for a different cobot. Therefore, the picker assumes a supportive role of fulfilling a task that the cobot is unable to perform (picking an item). In contrast, when the picker has a leading/active role, she navigates from one location to the other, performs the pick tasks fed to her via communication devices (e.g. pick by voice, pick to light, etc.) and drops off the retrieved items on the robot that plays the supportive role of following the picker in close proximity.

3. Literature and hypotheses development

3.1 Collaborative productivity, collaborative accuracy and human pick time

In human–robot collaborative order picking, productivity is influenced by both the collaborative and the non-collaborative components of the task. In both leading/following

collaborative order picking configurations, the collaborative portion of the task consists of humans and robots interacting, while the non-collaborative portion takes place when the robot carries completed orders autonomously to the depot station. Relieving pickers from unnecessary travel allows for greater utilization of the human workers, and, as such, recent research argues that it may translate to a higher system productivity of cobotic picking solutions when compared to traditional order picking (Azadeh *et al.*, 2025; Ghorashi Khalilabadi *et al.*, 2025; Löffler *et al.*, 2018).

System productivity at a macroscopic level typically denotes the correctly completed tasks per unit of time and, therefore, in this order picking context, the number of correctly picked orders per unit of time (Figure 1). Even though the time saved from a reduction of unnecessary traveling (the non-collaborative portion of task) is, by default, beneficial for system productivity, it is not currently clear how collaborating with robots will affect the bulk of an order picker’s job. Although human travel is reduced, the job still requires traveling, searching, picking and other minor activities (Tompkins *et al.*, 2010). The performance on these tasks is now a product of HRC. *Collaborative productivity* at the microscopic level of the order picking task denotes the portion of system productivity that contains the collaborative effort of picker and robot while they are paired (traveling together, being stationary together or interacting). This may be measured as the number of correctly picked number of orders (or order lines in general) per unit of time *during* the collaborative process and is inevitably capped by the productivity of the slower partner or the system bottleneck.

Assuming that robot speed is capped to warrant safety close to humans and is standardized across different setups, the spatial configuration of the leader–follower setups inevitably influences the output of the system. As such, when the human is leading (robot following/supporting) and has a higher degree of autonomy, she determines her own work pace, which leads to a larger variation in productivity. In contrast, when the robot leads, the picker’s pace may be constrained by the robot predefined speed, which now becomes the bottleneck. From a Self-Determination Theory (SDT) perspective (Deci and Ryan, 2012; Deci *et al.*, 2017), human motivation and performance are fundamentally shaped by the satisfaction of three basic psychological needs: autonomy (the need to feel volitional and self-directing), competence (the need to feel effective and capable) and relatedness (the need for meaningful connection with partners). Research demonstrates that when these needs are satisfied, individuals exhibit autonomous motivation, which reliably predicts enhanced performance, persistence and engagement across diverse work settings. The human-leading and robot-following configurations create distinctly different psychological environments that either support or constrain these fundamental needs. In human-leading setups, workers exercise relatively high (perceived) decision autonomy over navigation and pacing – directly satisfying the autonomy need. This autonomy support fosters autonomous motivation, which is associated with higher productivity compared to controlled motivation (Hagger *et al.*, 2014). Research in organizational psychology demonstrates that autonomous motivation leads to greater

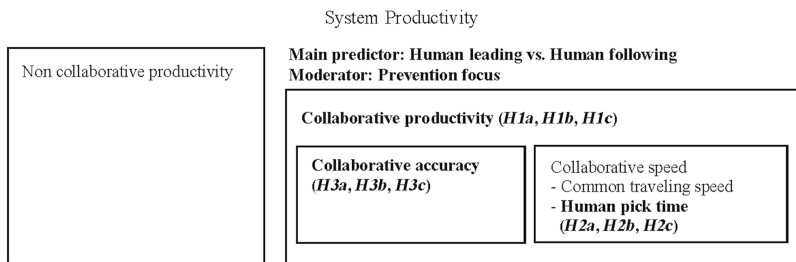


Figure 1. Components of HRC system productivity. Bold type = investigated performance indicators. The numbering of hypotheses is consistent with the sequence of the models tested in Section 4

persistence, concentration and effort, as well as enhanced proactive problem-solving behaviors. Therefore, both from the perspective of the structural dynamics of the setup and the perspective of behavioral theory, we expect that a setup with a *human leading* a robot is always equal or better than a setup with *human following* a robot in terms of collaborative order picking productivity.

H1a. Human leading the robot outperforms human following the robot in terms of collaborative order picking productivity.

To further dissect the human–robot collaborative productivity, we reduce it to the two key components of collaborative speed (or throughput) and collaborative accuracy (Figure 1). Collaborative speed simultaneously depends on the speed and capabilities of both robots and humans. Even though robot moving speed can be designed or assessed with great accuracy (Azadeh *et al.*, 2019), the speed of human order pickers varies across individuals (e.g. De Vries *et al.*, 2016a, 2016b) and consists of processing speeds of several sub-tasks (walking, searching, picking, etc.). Recognizing that the common movement speed is capped by the slower partner and equal across both scenarios, we can only differentiate across the two setups in the time spanning from the joint arrival to the joint departure from the pick location. While both partners are stationary at the pick location, the time spent there includes the time required by the human to complete the search for the required item and execute the pick. As such, the speed at which the human performs these actions may be directly attributed to her (denoted hereafter as *human pick time*). When the human is leading (*robot following*), the human has to assume a fully active state, process the order picking information and instructions at the beginning of her order picking cycle to initiate the movement, search until she reaches the pick location and execute the pick promptly after arriving at the pick location. In comparison, when the human is following (*robot leading*), the human has more choice on the level of involvement with the task across a continuum from being fully active to fully passive. On one extreme, she can be fully active: process the picking information and search before reaching the pick location and be prepared to execute the pick as soon as she reaches the pick location. On the other extreme, she can be fully passive: process the picking information and search for the pick only after the robot stops at the indicated picking position, then execute the pick. Therefore, human pick time when the human is leading will always be equal to or better (shorter) compared to when the human is following:

H2a. Human leading the robot outperforms human following the robot in terms of human pick time.

Conversely, collaborative order picking accuracy is expected to be higher when the robot is leading (*human following*), as the robot decreases the opportunity for mistakes by physically stopping at the correct pick location, limiting the human picker's choice. The robot's precise stopping at mapped locations and presentation of item information creates what SDT identifies as informational feedback – clear, immediate cues about correct actions that could significantly reduce errors (Deci *et al.*, 2017). The robot's structured guidance functions as an external error-prevention system, offering environmental support that compensates for human fallibility in complex pick environments. This represents the competence-support component of SDT, where clear feedback and environmental structure enhance task accuracy even when autonomy is somewhat constrained. This is also supported by the Leader–Member Exchange (LMX) Theory (Schriesheim *et al.*, 1999), which focuses on the quality of the relationship between leaders and followers. Applied to robots leading with high-quality LMX (e.g. when robots give clear instructions and allow some interaction), it may lead to greater levels of compliance of the humans (see Gerstner and Day, 1997; Tsai *et al.*, 2022). Together, this leads to the following hypothesis.

H3a. Human following the robot outperforms human leading the robot in terms of collaborative order picking accuracy.

3.2 Role of prevention focus in HRC

While the structural dynamics and behavioral context of HRC affect performance outcomes, individual human characteristics also play a key role in determining productivity and accuracy. Indeed, having the opportunity to work faster or more accurately does not necessitate actually doing it. Prior research has highlighted the importance of differences in worker personality characteristics in affecting order picking performance in both manual (e.g. [De Vries et al., 2016a](#)) and technology-assisted settings (e.g. [De Vries et al., 2016b](#)), as well as for a multitude of human factor aspects ([Grosse et al., 2017](#)).

Despite the growing adoption of HRC in order picking, behavioral research on the topic remains limited. Much of the existing literature on human–robot collaborative order picking relies on simulated settings (e.g. safety: [Inam et al., 2018](#); movement intention: [Petković et al., 2019](#)). These simulations are subject to assumptions that are usually not empirically validated. Even studies that utilize empirical data to assess theoretical models on scheduling coordination acknowledge their limitations to fully capture actual human working behaviors ([Wang et al., 2022](#)). Empirical findings suggest that HRC can enhance picker job satisfaction and self-efficacy and, under certain conditions, improve self-esteem ([Pasparakis et al., 2023](#)). Additionally, some evidence suggests that higher HRC workload reduces boredom but may impair performance when autonomy is low ([Hosseini et al., 2024](#)).

To better conceptualize human responses to the collaboration with robots, we consider the concept of prevention focus. Specifically, we theorize how the imposed cooperation in an interdependent setup evokes a behavioral response that is dictated by a prevention focus and affects task engagement.

In HRC picking environments, the final outcome is a sequence of smaller tasks with a common end-goal, performed alternately by the human and the robot (with the occasional overlap). In such interdependent settings, the picker can only achieve maximum productivity by synergizing with the robot and is effectively encouraged to work in cooperation with the robot (and not in competition with it) ([Rosenbaum et al., 1980](#)). The introduction of robots forces the picker into collaboration with a fixed-resource teammate, resulting in task interdependence constrained by the capabilities of the robot ([Steiner, 1972](#)). As a result, HRC can be perceived as inherently threatening to worker productivity and the subsequent earning potential. This might be especially salient for workers who are naturally inclined to strategically devise their behaviors to avoid negative outcomes or “high prevention focused” individuals, according to regulatory focus theory ([Higgins, 1998](#); [Wallace and Chen, 2006](#)). Such individuals are expected to be more engaged within a collaborative group setup and are consequently likely to have the intention to excel at their tasks, despite any imposed limitations ([Beuk and Basadur, 2016](#)). Prevention focus describes this inclination of a risk-minimizing individual, and it may additionally enhance the effects of cooperative setups on the performance of partners (e.g. [Beersma et al., 2013](#); [Chernikova et al., 2017](#)). In the context of warehouse order picking, [De Vries et al. \(2016a\)](#) have shown that prevention focus explains an important part of the productivity variation in sequential (interdependent) order picking assignments under cooperation setups. We, therefore, hypothesize that:

- H1b.* A higher human prevention focus leads to higher collaborative productivity in human–robot collaborative order picking.
- H2b.* A higher human prevention focus leads to a shorter human pick time in human–robot collaborative order picking.

Building further on the point that a strong prevention focus will lead to more vigilance and mistake-avoidance efforts ([Higgins, 1998](#); [Wallace and Chen, 2006](#)), we additionally hypothesize that:

- H3b.* A higher human prevention focus increases collaborative order picking accuracy in human–robot collaborative order picking.

Beyond the overall effect of prevention focus, we expect that its impact might vary between the scenarios of *human leading* and *human following*. Prior research has shown that human order picking productivity under cooperation-based setups declines as the degree of interdependence between (human) workers increases (De Vries *et al.*, 2016a). This reduction in productivity when partners collaborate in sequential settings is often attributed to work-rate adjustment, influenced by the observed productivity of the partner (Cantor and Jin, 2019; Doerr *et al.*, 2002; Powell and Schultz, 2004; Schultz *et al.*, 1998).

Such state-dependent behavior may also occur in the *human-following* scenario, which resembles a cooperation-based sequential setting with the human tasks contained in-between robot tasks, and the robots playing the role of downstream partners (i.e. in front of the human worker). In such a setting, humans may choose to either slow down their task to match the speed of their robotic partner or instead focus on maximizing collaborative productivity, for example, by preparing for the upcoming tasks before reaching the pick location. As individuals with a high prevention focus experience the threat that robots pose on their productivity (and subsequent rewards) more saliently, they are more likely to adopt strategies that mitigate their potential losses. Consequently, differences in human behavior under the *human-following* scenario may partially account for the productivity gap between the scenarios (according to H1a). We therefore theorize that pickers with a higher prevention focus will maintain their productivity efforts even when following robots:

H1c. Human prevention focus moderates the effect of human leading vs. human following on collaborative productivity. A higher prevention focus diminishes the gap in collaborative productivity across these conditions.

H2c. Human prevention focus moderates the effect of human leading vs human following on human pick time. A higher prevention focus diminishes the gap in human pick time across these conditions.

Additionally, the positive effect of prevention focus on collaborative accuracy is more likely to occur in the *human leading* scenario, where, according to H3a, there is an additional risk for mistakes. Therefore, we hypothesize that:

H3c. Human prevention focus moderates the effect of human leading vs human following on collaborative accuracy. A higher prevention focus diminishes the gap in collaborative accuracy across these conditions.

4. Methodology

4.1 Participants and research setting

To empirically test our hypotheses and investigate the factors that govern HRC, we designed and conducted a large-scale controlled physical experiment involving 60 participants. This methodology is well-suited for empirically investigating the effects of human behavior in group dynamics (Bendoly *et al.*, 2010). This setup required substantial effort due to its one-on-one nature, with each participant engaging in a three-hour session under the supervision of two experimenters (one for robot supervision, one for quality control and restocking).

To ensure sufficient statistical power to test our hypotheses, we used a within-subjects design, exposing all 60 participants to both setups of *human leading* and *human following* (procedure and randomization explained further under §3.2). This enabled us to control for more individual extraneous variables and made it easier to detect differences in performance across treatments (Eckerd *et al.*, 2021).

The participants' pool consisted of Dutch students at a vocational shipping and transport college in the Netherlands, training to become future logistics employees. Using vocational students promotes practical generalizability and relevance for our findings, in accordance with previous studies, which suggest that vocational students are more appropriate subjects for

physical order picking experiments compared to the more common practice of deploying university students (De Vries *et al.*, 2016b). Participants were recruited randomly from a pool of interested senior students after the experiment was advertised as an attractive opportunity to gain experience in working with cutting-edge technology. The proportion of female to male recruited students (10:50) was representative of the college’s student population. The average age of the participants was 18.0 years ($sd = 1.3$ years). To motivate sustained effort for the session, the participants were compensated with the combination of three components that reflect reward systems observed in practice: (1) a certificate of training experience (emulating work experience), (2) a flat monetary reward for participation (base compensation/wage), as well as (3) additional bonus monetary prizes for the top three performers that picked the largest amount of correct order lines (sum of all monitored rounds), among all participants (performance bonus). The combination of a representative research sample and a realistic reward system allowed us to obtain findings that are methodologically rigorous, yet meaningfully generalizable to practice.

To accurately replicate the order picking process, we erected a dedicated experimental warehouse in the Netherlands with 300 distinct pick locations (Figure 2). The setup was designed to resemble a typical shelf warehouse used worldwide for storing small items in small quantities. As one of Europe’s largest logistics hubs with a considerable number of implementations of warehouse cobotics over the last years (Globe Newswire, 2025), the Netherlands is an appropriate research context for the experiment. The warehouse was constructed in a secluded building, with professional metal racks in a layout of two and a half aisles, consisting of ten sections per aisle. Each rack was 2,100 mm high and was divided into four levels (vertically) and into three pick positions per level (horizontally). Pick locations were systematically labeled marking the aisle-section-level-position identity in natural progression (e.g. an item at the pick location A.04.02.03 is in aisle A, in the fourth section, on the second level and at position three). Each pick location hosted a unique household product (volume from 0.1 to 5 l, and weight from 10 to 2000 g), with enough stock to allow for

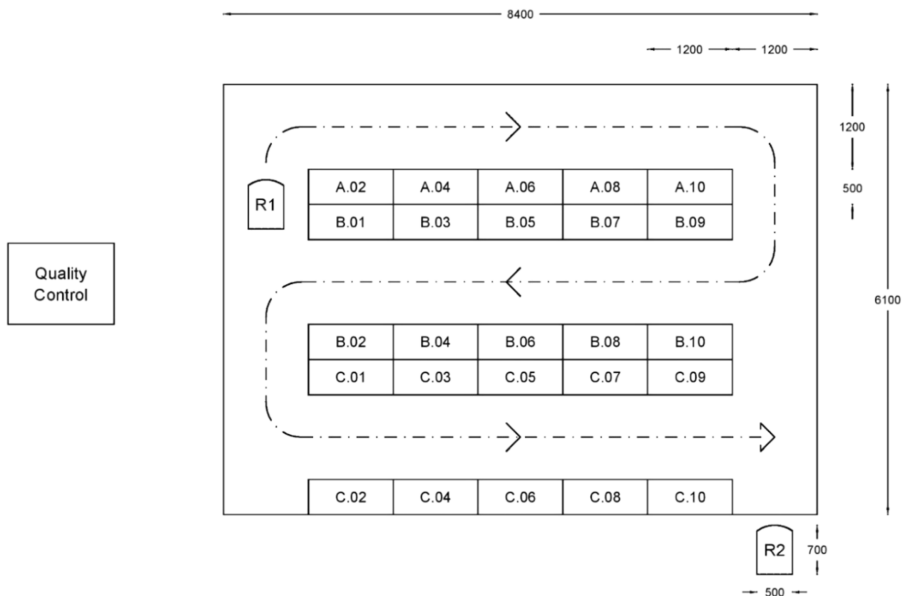


Figure 2. Warehouse layout (units in mm)

uninterrupted picking during each picking round. The realistic warehouse design, the order process and the robot collaboration contribute to enhancing the external validity of our results.

The participants were instructed to perform a simplified order picking task that can be summarized in four steps:

- (1) Locate the next requested product on the order list.
- (2) Pick the specified quantity and place them in the order crate.
- (3) Confirm the pick via the touchscreen interface.
- (4) Repeat the process with the next order line until the order is complete, then proceed to the next available order.

Orders were sequentially fed to the participants via a simple touchscreen interface and were virtually inexhaustible, ensuring that participants were engaged at all times during the pick runs. The interface provided information (location code, name of the product, number of requested units) for all the order lines of the order at hand. To provide enough variation in the task, three large sets of random orders were generated, with each order consisting of 5 to 10 order lines (5: 40%, 6: 25%, 7: 15%, 7: 10%, 9: 5%, 10: 5%) requesting a unique product. Each of the 300 pick locations had the same probability of being selected, with the only constraint that every order included at least one item from each of the three aisles, so as to force pickers to visit all three aisles per order. To further standardize the movement, we implemented an alternating sequence of visiting the aisles (A-B-C for orders with odd numbers and C-B-A for orders with even numbers). As such, we nullified the opportunity for taking shortcuts and created a continuous S-shaped movement for the picker, requiring the same total travel distance to fulfil each order. The number of units to be picked for each order line ranged from one to three (one: 50%, two: 30%, three: 20%).

4.2 Experimental procedure

The experiment consisted of two phases. Phase 1 commenced at the participant's arrival at the warehouse. First, they filled out a participation consent form along with a questionnaire (demographics, regulatory focus) and watched a series of videos explaining the procedure of the experiment and the general rules of order picking. All forms, questions and videos were in Dutch (the native language of all participants). Second, they performed one round of manual order picking using a pick trolley. This allowed the participants to get familiar with the warehouse and the mechanics of order picking, such as objectives, routing logic, naming schemes, electronic pick lists and touchscreen interfaces.

In Phase 2, participants were introduced to the manipulation following a within-subjects design. Each participant was assigned to a (counterbalanced) sequence of the two collaborative order picking conditions, of (1) *human leading* and robot following, (2) robot leading and *human following*. In total, every subject participated in three 20-min monitored order picking rounds. Three distinct order sets were also randomly assigned to each of the three rounds, resulting in a grid of $2 \times 3! = 12$ different (*scenario*) \times (*order set sequence*) assignments. Figure 3 depicts the participant's journey. Before each round, the participants followed instructions and practiced the task, and after each round, they filled out a questionnaire which captured their emotional states as well as their perception of the performed task. The monitored subjective/perceptual constructs are not affecting or overlapping with any objective outcomes of the current study and were used in follow-up research (Pasparakis *et al.*, 2023).

4.3 Manipulation

To implement the manipulation, we deployed two identical and undistinguishable prototype autonomous robots with localization accuracy of ± 20 mm accuracy facilitated by a 270-degree laser scanner. Each robot was fitted with a touchscreen mount and enough space for

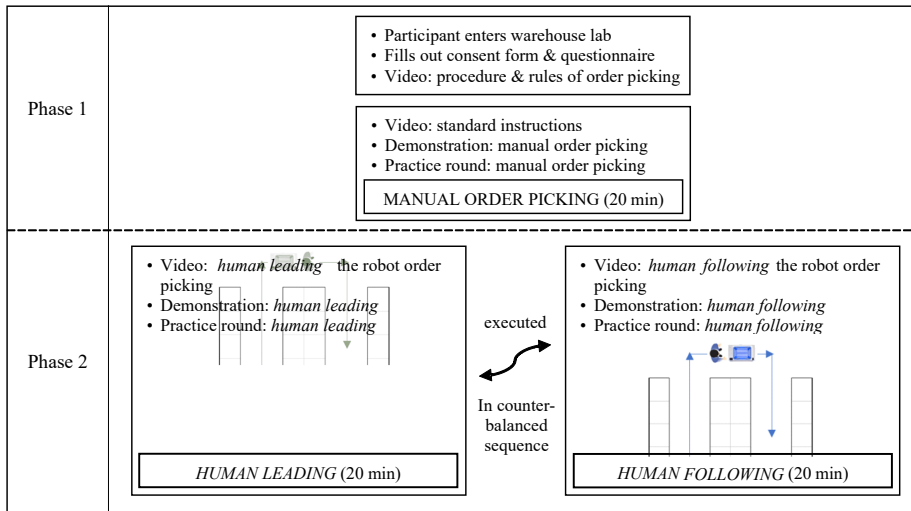


Figure 3. Participant's experimental journey

transporting one empty order crate. The size of the orders and the speed of unloading were set such that a human picker never had to wait for a robot for the next order to be picked.

In the *human leading* condition, a small touch screen was fitted on the subject's non-dominant arm. The subject was explicitly informed that she was in control and leading the process, while the robot was there to support by carrying the crate for her. The participant used the touch screen to locate the next pick location, and the robot followed her movement at a close distance, carrying the order's crate. Conversely, in the *human following* setting, a larger touch screen was fitted on the robot's back to account for the greater viewing. In this scenario, the robot was framed to be in control and leading the process, while the subject was there to support it by following and putting the products in the crate. In this scenario, the robot navigated to the designated pick location, while the subject followed it to perform the pick at the destination.

The primary distinctions between the two scenarios were characterized by: (1) enforcing the leader-follower collaboration dynamics due to the relative spatial positioning of the two collaborators and (2) verbal framing of the leading (active) vs following (supportive) roles of the collaborators (Figure 4).

Although participants were led to believe that the two scenarios were fundamentally different, they were designed to ensure that the picking performance was affected by factors only attributable to the participants. To establish comparability between setups and eliminate the confounding variable of robot performance, we employed a Wizard of Oz approach. Participants were informed that the robots were intelligent enough to locate the products in the warehouse or to follow the human autonomously. In reality, the robots' movements for each set of orders were predefined and identically carried out at the same speed across scenarios for movements between the same locations, with only variation to accommodate human positioning. This offset ensured that the picker positioned herself always in the midpoint of the section where the requested product was stored, and, therefore, standardized stopping nodes when picking the same product across scenarios. Naturally, two sections located on either side of the same aisle share a common stopping node (e.g. B.03 and B.04), and each aisle or half aisle has five unique stopping nodes.

The sequence of commands was preprogrammed in the robots, with navigation instructions triggered each time the participant confirmed the previous pick with a tap on the screen. In the



Figure 4. HRC settings. Left: Human leading and robot following. Right: Human following and robot leading

robot leading condition, the picker's movement did not affect that of the robot. When the picker was leading, the robot followed the predefined route without overtaking the picker, but stopped at a safe distance behind the picker, creating the illusion of following. As such, it was not possible for participants to notice the difference, and we did not notice any signs that indicated participants observed this. To maintain a continuous workflow, the finishing point of all orders was positioned in the immediate proximity of the next order's starting point (due to the imposed S-shaped route), meaning that the participant had direct access to the second robot loaded with the next order. While the human was occupied with fulfilling the next order, the first robot transported the crate to the quality control station without interfering with the picker's or the second robot's movements. As such, unnecessary traveling was eliminated, and the humans were continuously occupied with tasks that fall under the collaborative performance denotation (Figure 1). As the robot movement and speed were the same across scenarios and for all participants, all variation in the observed outcomes is a result of variation in human behavior and performance.

Throughout both conditions, the subject retained continuous access to order information on the screen interface. A subtle but crucial difference between the scenarios was that in the robot-leading scenario, the subject could decide *when* to retrieve the information from the screen. Therefore, the subject could adopt a stance ranging from active (processing the information at the beginning of the pick cycle) to passive (processing the information after reaching the pick location). In both scenarios, the robot automatically transported the crate to the quality control station (Figure 2), as soon as the subject completed the order. The subject then interacted with

the next robot, which was fitted with an empty crate for the next order. The process was repeated until the predefined time was over.

4.4 Operationalizing objective performance outcomes (dependent variables)

Throughout both HRC rounds, we closely monitored and assessed the performance of the participants in terms of collaborative productivity, collaborative accuracy and human pick time. We assume that participants became proficient with the order picking method during the training and practice phases of each round.

Collaborative productivity was quantified as the number of correctly completed order lines during each 20-min picking run. In practice, incorrect lines must be repaired and lead to a net loss of throughput, as pointed out by previous literature on order picking experiments (e.g. [De Vries et al., 2016a](#)). Nevertheless, we performed robustness checks in our final models using the total number of completed order lines, ensuring that the established relationships are analogous. In the HRC scenarios, order picking throughput is capped by the performance characteristics of the robot. Therefore, the aggregate number of items picked reflects the collaborative productivity or the productivity that the subject achieved, reflecting any hindrance imposed by the robot.

To further assess human engagement within the collaborative setting, we monitored the human pick time, defined as the duration that subjects took from the moment they reached the pick location until the moment they completed (confirmed) the pick. We retrieved this information from the interface and robot movement logs, which were recorded independently from the quality control station data. We additionally compared the number of confirmed picks (from the robot logs) with the total number of picks (recorded at the quality control station), to cross-validate the accuracy of the records kept by the experimenters. It should be noted that the human pick times were segmented into three clusters: 2,720 picks of products located in the same aisle but at different stopping nodes (26 combinations of stopping nodes), 3,609 picks of products located in different aisles (46 combinations of stopping nodes), and 1,165 picks of items located at the same stopping node (10 distinct stopping nodes). This segmentation serves as an internal robustness check on the results produced by the analysis of the aggregate/quality control data across all potential types of node-to-node movement. As such, the split allows for a more appropriate fit of the final models by individually capturing the slightly different spreads among the respective human pick time distributions. Cluster III (C-III) is especially interesting, as it differs distinctly from the other two: the picker does not need to travel during the pick cycle (the robot also stands still), which means that the human pick time can be attributed entirely to the human worker. Note, we only have time stamps of the actual picks; the time between two picks includes traveling, searching and picking. Consequently, the time pickers actually spend picking is a derived metric, which we construct based on the known travel speed of the robot. However, in practice, this speed might slightly deviate, for instance, when a picker blocks the path of the robot. The likelihood of these deviations occurring might differ between our experimental conditions. We are primarily interested in differences in the *Human* part of the pick time (not the joint human–robot pick time, which may include some waiting of the human for the robot or vice versa), as induced by differences in the setup. To find differences in human pick time not induced by the setup itself, but by how humans respond to it, it is therefore interesting to look at settings which are completely identical between the two conditions. These can be found when multiple items need to be picked from one location (Cluster III), as there is no robot-induced waiting here.

Collaborative accuracy was operationalized in terms of *inaccuracy* by counting the number of mistakes made during the 20-min run, including both location errors (wrong item picked) and quantity errors (wrong number of units of correct product picked).

4.5 Operationalization of independent and control variables

Our core independent variable, the human leading vs. human following manipulation, was introduced as a dummy variable with the “human following” condition as the reference level. Our moderator, *prevention focus*, was measured using a six-item structure (Table 1) from the Regulatory Focus at Work scale proposed by Wallace and Chen (2006). This well-established scale holds its internal consistency and validity in various contexts, including warehouses (De Vries et al., 2016a, c). Additionally, to account for individual differences, we measured each participant’s initial performance during the introductory manual pick run. As the participants developed an unimpeded natural pace during this run, the number of correct order lines picked reflects their natural performance level in order picking operationalized as a *base level of performance*. This baseline was used to account for inter-individual variance across the HRC conditions.

Finally, we collected demographic data, including *age*, *sex*, *education level* and *order picking experience* of the participants. Only the *order picking experience* was included in the final reported models due to limited variance across subjects in *age* and *education levels* and the absence of a significant relationship found for *sex*.

5. Results

5.1 Descriptive statistics

We first establish an overview of all monitored variables by analyzing the descriptive statistics. Table 2 presents means, standard deviations and bivariate Pearson correlations of relevant subject-level variables that are used in the final models. *Prevention focus* ($\alpha = 0.80$) is more

Table 1. Regulatory focus at work items measuring prevention focus (Wallace and Chen, 2006; 5-point Likert scale)

Please indicate how often you think of the following issues while working	
1	Following rules and regulations
2	Completing work tasks correctly
3	Doing my duty at work
4	My work responsibilities
5	Fulfilling my work obligations
6	On the details of my work

Table 2. Descriptive statistics and bivariate Pearson correlations for relevant subject-level variables

Variables	Mean	Sd	1	2	3	4	5	6
<i>Independent</i>								
1 Order picking experience (m)	2.12	5.40						
2 Prevention focus	4.24	0.48	0.04					
<i>Dependent: Human leading</i>								
3 Collaborative productivity (correct order lines/20 min)	67.90	5.51	-0.13	0.04				
4 Inaccuracy (errors/20 min)	0.88	1.03	-0.24	0.04	-0.21			
5 Average human pick time (sec)	3.18	0.80	0.07	-0.10				
<i>Dependent: Human following</i>								
6 Collaborative productivity (correct order lines/20 min)	62.85	5.20	-0.06	0.20				
7 Inaccuracy (errors/20 min)	0.70	1.00	-0.05	-0.13				-0.46
8 Average human pick time (sec)	3.41	1.18	0.13	-0.22				

Note(s): N = 60. italic values indicate a significant relationship, $p < 0.05$

strongly correlated with *collaborative productivity* in the *human following* scenario than in the *human leading* scenario.

5.2 Human–robot collaborative order picking productivity

To predict overall human–robot *collaborative productivity*, we fit linear mixed-effects models using the package “lme4” (Bates *et al.*, 2007) in R version 4.0.3 (R Core Team, 2020). We control for the *order picking experience* and *base performance* of the pickers to explain part of the variance in the dependent variables. *Base performance* is modeled as an ordinal variable {low, medium, high}, calculated by ranking the pickers according to their *total productivity* in the initial manual round and by splitting them into three equal groups. This transformation allowed for consistent application across all models, reducing degrees of freedom and ensuring model convergence. Furthermore, we control for the specific *order set* used in each pick round to capture any differences in performance due to the different pick lists. Given the nested nature of the data at the picker level across the two scenarios (*human leading* the robot vs *human following* the robot), we introduce a random intercept to account for subject-level variation. Scenario conditions are introduced as a binary variable, with a value “1” for the *human leading* the robot and a value “0” for the *human following* the robot. To address potential spill-over effects of learning or fatigue, we additionally control for the *sequence* of the scenarios.

We first test the objective differences between scenarios (*scenario models*), before introducing behavioral components in subsequent steps. Picker *prevention focus* is tested both as a standalone variable (*behavioral models*) and in interaction with the scenario (*interaction models*). Model improvements were assessed using ANOVA tests (West *et al.*, 2014), with statistically significant model improvements indicated on the respective tables with the estimation criterion in bold type. The final reported models are fitted using the restricted maximum likelihood ratio parameter estimation criterion, which reduces bias in variance components compared to maximum likelihood ratio tests (Wu *et al.*, 2017).

Table 3 presents the results of the linear mixed-effects models predicting total human–robot order picking *collaborative productivity*. M1a indicates that *human leading* the robot positively predicts *collaborative productivity*, supporting Hypothesis H1a. However, we find no sufficient evidence in support of H1b nor H1c, as M1b and M1c show no significant improvement on M1a. We additionally used the “Emmeans” package (Lenth, 2018) in R version 4.0.3 (R Core Team, 2020) to conduct an estimated marginal means analysis on M1a, which revealed a +8.3% average productivity advantage of *human leading* compared to *human following* scenarios (66.2 vs 61.1 picks, using average levels of all control variables). The importance of the behavioral component is evident due to the significant main effects of *prevention focus* ($p < 0.05$) when introducing interaction effects of *prevention focus* and *scenario* ($p = 0.14$) under M1c. As the aggregate data are bound by the performance of the robot and do not allow for large variance in performance among participants, we further investigate the behavioral aspect of HRC on the more refined level of analysis with data on *human pick time* in Section 4.3. In addition, we find that *base performance* in manual order picking is transferable to human–robot collaborative order picking, positively predicts *collaborative productivity*, as the more competent participants continue to outperform the less competent ones even after the robot introduction. The *sequence* of the scenarios is not statistically significant; therefore, there is no evidence of a spill-over effect between order picking cycles. The different *order sets* appear to influence *collaborative productivity*. To investigate the underlying reason, we performed a post-hoc analysis on the dispersion of products requested. Both the mean nodes-to-visit per order for each *order set* and the units-to-pick per order line are similar in most *order sets*. We, therefore, attribute any differences across *order sets* to differences in the physical characteristics of the requested products (e.g. more complicated products, stacked products).

Table 3. Linear mixed-effects models predicting collaborative productivity in human–robot collaborative order picking

Independent variable	Dependent variable: Collaborative productivity (correct total order lines picked in 20 min)					
	Scenarios model (M1a)		Behavioral model (M1b)		Interaction model (M1c)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	60.213	1.048	54.201	4.199	50.895	4.763
Base performance: Medium	4.472	1.105	4.828	1.120	4.841	1.119
Base performance: High	7.935	1.085	8.052	1.077	8.066	1.076
Order picking experience	−0.117	0.086	−0.126	0.085	−0.125	0.085
Order set #2	−2.301	0.668	−2.229	0.669	−2.080	0.671
Order set #3	−3.766	0.654	−3.660	0.657	−3.620	0.652
Scenario sequence	1.598	0.905	1.444	0.902	1.441	0.901
Scenario: <i>Human leading</i>	5.074	0.490	5.074	0.490	11.453	4.381
Prevention focus (Scenario: <i>Human leading</i>) *			1.388	0.939	2.150	1.073
(Prevention focus)					−1.504	1.026
Restricted maximum likelihood criterion	634.0		630.2		622.8	
Number of subjects	60		60		60	
Number of rounds	120		120		120	

Note(s): Italic type = significant at $p < 0.05$

5.3 Human pick time

To analyze the human engagement and intention for productivity, we extracted the *human pick times* from 7,494 individual picks (mean = 124.9 picks/subject) using robot navigation and interface logs. Pick time was calculated by subtracting travel time from the total order pick cycle time. This number includes instances of moving from a previous product location to the current location and therefore excludes the first pick of every order, which might include additional setup time-costs or time-savings by preprocessing information before initiating the order picking cycle. Furthermore, we exclude erroneous entries due to network disconnections or other system malfunctions that happened at random times during the experiment (representing ~3.7% of picks).

As explained in the methodology section, human pick time was split into three separate clusters. Across all three clusters, the *human pick times* were left-bounded by zero and positively skewed. To address this, we re-expressed the times using the x^{λ} function, following the Tukey ladder of Powers (Tukey, 1977). We calculated the lambdas that maximize the Shapiro–Wilks W statistic for the distributions of each of the three clusters, using the package “rcompanion” (Mangiafico, 2020) in R version 4.0.3 (R Core Team, 2020). The calculated lambdas for the three clusters were positive {C-I: 0.325, C-II: 0.200, C-III: 0.675}.

To predict *human pick time*, we fit linear mixed-effects models on the merged timer data and survey data, using the package “lme4” (Bates et al., 2007) in R version 4.0.3 (R Core Team, 2020). The data are nested at both picker level and stopping node combination level, as pickers picked multiple times on multiple stopping node combinations (with different traveling times and characteristics that might influence *human pick time* as well). To address the cross-nested nature of our data (Bates et al., 2014), we introduce a random intercept to account for the variability of random subject and the effects of random stopping node combination. We also control for *order picking experience* and *base performance* to explain part of the variability in the dependent variables, as described in the previous models. Additionally, we control for *progress* in the pick round using the square root of the pick timestamp in seconds [0, 1,200]. We introduce this measure to capture a combination of effects

that might take place while progressing through the pick run, such as flow, learning and fatigue. The final models are fitted using the restricted maximum likelihood ratio parameter estimation criterion.

5.3.1 *Human pick time: same aisle, different stopping node.* Table 4 presents the results of the linear mixed-effects model predicting *human pick time* for items located in the same aisle (C-I). The best fitting model (M2c,i) shows that the direct and interaction effect of *prevention focus* and *scenario* are statistically significant, providing support for Hypotheses 2a, 2b and 2c. As such, picker *prevention focus* moderates the effect of picker leading vs following the robot on picking productivity. A two-way interaction plot of the observed range of picker *prevention focus* predicting *human pick time* re-transformed to seconds (Figure 5), demonstrates that high prevention-focused individuals effectively bridge the productivity gap between the scenarios. Further marginal means analysis indicates an average productivity advantage of 40.1% ($\Delta = 1.23, p < 0.05$) for high prevention-focused pickers over low-prevention focus ones in the *human following* scenario (Figure 5). These results were calculated on average over all control variables using “Emmeans” (Lenth, 2018) in R version 4.0.3 (R Core Team, 2020). Additionally, results confirm that *base performance* and *progress* both predict shorter *human pick times*.

5.3.2 *Human pick time: different aisle.* Table 5 presents the results of the linear mixed-effects model predicting *human pick time* for items located in different aisles (C-II). The best fitting model (M2b,ii) includes the main effects of *human leading* and *prevention focus*. *Human leading* and *prevention focus* are both beneficial for shorter *human pick times*, therefore supporting Hypotheses 2a and 2b. Interestingly, only high competence is transferable from manual picking to human–robot collaborative order picking, while prior *order picking experience* is detrimental to *human pick times*, though with a small effect size.

5.3.3 *Human pick time: same stopping node.* The models predicting *human pick time* at the same stopping node (C-III) violated the assumption of homoscedasticity (Levene’s test) and

Table 4. Linear mixed-effects models predicting human pick time of products located in same aisle (different stopping node)

Independent variable	Dependent variable: Human pick time in same aisle (Tukey transformation, $\lambda = 0.325$)					
	Scenarios model (M2a,i)		Behavioral model (M2b,i)		Interaction model (M2c,i)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	<i>1.497</i>	0.049	<i>1.806</i>	0.152	<i>1.976</i>	0.170
Base performance: Medium	<i>-0.152</i>	0.041	<i>-0.170</i>	0.040	<i>-0.170</i>	0.040
Base performance: High	<i>-0.269</i>	0.040	<i>-0.274</i>	0.039	<i>-0.274</i>	0.039
Order picking experience	0.005	0.003	0.005	0.003	0.005	0.003
Scenario sequence	<i>-0.015</i>	0.033	<i>-0.007</i>	0.032	<i>-0.007</i>	0.032
Progress (square root of pick timestamp)	<i>-0.006</i>	0.001	<i>-0.006</i>	0.001	<i>-0.006</i>	0.001
Scenario: <i>Human leading</i>	<i>-0.084</i>	0.017	<i>-0.084</i>	0.017	<i>-0.409</i>	0.145
Prevention focus			<i>-0.072</i>	0.034	<i>-0.112</i>	0.038
(Scenario: <i>Human leading</i>) * (Prevention focus)					<i>0.076</i>	0.034
Restricted maximum likelihood criterion	3161.6		3162.0		3161.9	
Number of subjects	60		60		60	
Number of stopping node combinations	26		26		26	
Number of picks	2,720		2,720		2,720	

Note(s): *Italic type* = significant at $p < 0.05$

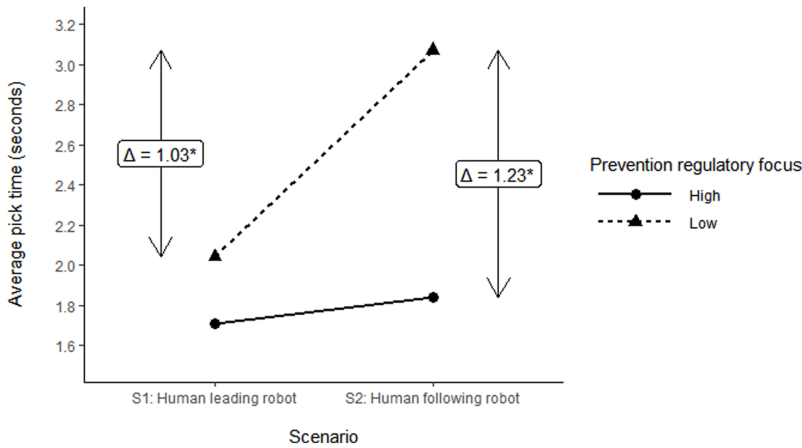


Figure 5. Two-way interaction plot between human leading vs human following scenarios and prevention focus, predicting human pick time on the same aisle, different stopping node. *significant difference in post-hoc test $p < 0.05$

Table 5. Linear mixed-effects models predicting human pick time of products located in different aisle

Independent variable	Dependent variable: Human pick time in different aisles (Tukey transformation, $\lambda = 0.200$)					
	Scenarios model (M2a,ii)		Behavioral model (M2b,ii)		Interaction model (M2c,ii)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	<i>1.087</i>	0.024	<i>1.269</i>	0.074	<i>1.263</i>	0.086
Base performance: Medium	-0.029	0.020	-0.039	0.020	-0.039	0.020
Base performance: High	-0.078	0.020	-0.081	0.019	-0.081	0.019
Order picking experience	<i>0.004</i>	0.002	<i>0.004</i>	0.001	<i>0.004</i>	0.001
Scenario sequence	0.017	0.016	0.022	0.016	0.022	0.016
Progress (square root of pick timestamp)	0.000	0.001	0.000	0.001	0.000	0.001
Scenario: <i>Human leading</i>	-0.042	0.010	-0.042	0.010	-0.030	0.086
Prevention focus (Scenario: <i>Human leading</i>) *			-0.042	0.016	-0.041	0.019
(Prevention focus)					-0.003	0.020
Restricted maximum likelihood criterion	<i>1379.0</i>		<i>1378.8</i>		<i>1384.8</i>	
Number of subjects	60		60		60	
Number of stopping node combinations	46		46		46	
Number of picks	3609		3609		3609	

Note(s): Italic type = significant at $p < 0.05$

were subsequently estimated using robust linear mixed-effects models (Table 6). The calculated robustness weights for the residuals and the random effects are numerous and, therefore, not possible to report in detail. We find that competence is transferable from manual picking to on-spot picking while collaborating with a robot. As with picks from different aisles (C-II), *order picking experience* is detrimental to *human pick times*. *Progress* is significant,

Table 6. Robust linear mixed-effects models predicting human pick time of products located at the same stopping node

Independent variable	Dependent variable: Human pick time at same stopping node (Tukey transformation, $\lambda = 0.675$)					
	Scenarios model (M2a,iii)		Behavioral model (M2b,iii)		Interaction model (M2c,iii)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	4.513	0.339	5.098	0.593	5.523	0.710
Base performance: Medium	-0.641	0.140	-0.674	0.138	-0.672	0.138
Base performance: High	-0.938	0.136	-0.949	0.132	-0.951	0.132
Order picking experience	0.021	0.011	0.022	0.011	0.023	0.011
Scenario sequence	-0.118	0.113	-0.105	0.110	-0.105	0.110
Progress (square root of pick timestamp)	-0.040	0.005	-0.040	0.005	-0.040	0.005
Scenario: <i>Human leading</i>	-0.066	0.085	-0.070	0.085	-0.877	0.745
Prevention focus (Scenario: <i>Human leading</i>) * (Prevention focus)			-0.137	0.114	-0.237	0.147
Number of subjects	60		60		60	
Number of stopping nodes	10		10		10	
Number of picks	1,165		1,165		1,165	

Note(s): *Italic type* = significant at $p < 0.05$

supporting the notion that it captures the effects of learning and flow. Most importantly, we did not find a significant difference across different levels of *prevention focus* and between the *human leading* vs. *human following* conditions. This was to be expected, as there is limited room for leadership dynamics to make a difference when both the robot and picker are stationary while picking.

5.4 Human–robot collaborative order picking accuracy

To assess the accuracy in HRC order picking, we use the number of errors made per 20 min interval (including incorrect item selection and incorrect quantity picked). Given the count-based, heavily skewed (high probability of a low number of errors) nature of the data, we choose a negative binomial distribution to describe them as the aggregation of randomly occurring events. We deploy the same control variables as used in the models predicting *collaborative productivity*, but add the number of correct lines picked per 20 min to account for additional error opportunities due to larger pick volume. We fit the models using an extended generalized linear mixed-effects model builder (Skaug *et al.*, 2014) in R version 4.0.3 (R Core Team, 2020).

Table 7 shows the results of the negative binomial mixed-effects models predicting human–robot collaborative order picking *inaccuracy*. The best-fitting model (M3a) shows that *human leading* positively predicts committing errors (*inaccuracy*), or, the inverse, that the *human following* positively predicts accuracy, therefore supporting Hypothesis 3a. Post-hoc analysis (Post-hoc 1: reevaluating the model with outcome variables: 1. Errors in quantity, 2. Errors of product type) suggests that the additional errors committed when humans lead the robots are wrong quantity picks. Neither the main effects (M3b) nor the interaction effects (M3c) of *prevention focus* are significant. Therefore, Hypotheses 3b and 3c are not supported by evidence in the data.

5.5 Quantification of practical impact

To quantify real-world impact, we assume that the typical order picker works for six net hours per day. We further assume that the pickers have immediate access to a new unoccupied robot

Table 7. Negative binomial mixed-effects models predicting *inaccuracy* in human–robot collaborative order picking

Independent variable	Dependent variable: Inaccuracy (total errors per 20 min)					
	Scenarios model (M3a)		Behavioral model (M3b)		Interaction model (M3c)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	<i>4.088</i>	1.714	<i>4.206</i>	1.901	<i>5.161</i>	2.112
Base performance: Medium	−0.386	0.331	−0.403	0.352	−0.414	0.353
Base performance: High	0.316	0.362	0.303	0.372	0.283	0.373
Order picking experience	−0.050	0.036	−0.050	0.036	−0.053	0.036
Order set #2	0.210	0.311	0.210	0.311	0.182	0.314
Order set #3	0.540	0.300	0.537	0.301	0.549	0.301
Scenario sequence	−0.289	0.249	−0.288	0.249	−0.287	0.249
Correct total order lines picked in 20 min	<i>−0.075</i>	0.029	<i>−0.074</i>	0.029	<i>−0.071</i>	0.029
Scenario: <i>Human leading</i>	<i>0.662</i>	0.263	<i>0.657</i>	0.264	−1.341	1.960
Prevention focus (Scenario: <i>Human leading</i>) *			−0.038	0.268	−0.305	0.374
(Prevention focus)					0.472	0.459
Log-likelihood	−131.09		−131.08		−130.55	
Number of subjects	60		60		60	
Number of rounds	120		120		120	

Note(s): *Italic type* = significant at $p < 0.05$

as soon as they complete their previous order. Therefore, we can directly transfer the numerical effect size estimations for order picking productivity, accuracy and intention for productivity, obtained from the estimated marginal means analysis (using the average levels of all control variables) and the interpretation of the results in Tables 3, 4 and 7. We calculate three quantifiable advantages:

- (1) Overall productivity advantage of the *human leading* compared to *human following* setup amounts on average to an additional 8.3% or $\left(\frac{66.2 \text{ order lines}}{20 \text{ min}} - \frac{61.1 \text{ order lines}}{20 \text{ min}} \right) * \frac{60 \text{ min}}{20 \text{ min}} * 6 \text{ hours} = +91.8 \text{ order lines per order picker per day}$ (Table 3). This advantage translates to +51 small orders per order picker per day or +2.6 large orders per order picker per day.
- (2) Time savings for high prevention-focused individuals can amount on average to $40.1\% * 15\% * 60 \text{ min} * 6 \text{ hours} = 21.7 \text{ minutes per day}$ (Table 4). This assumes that 15% of order picking time is real pick time (Tompkins *et al.*, 2010) and can be attributed to shorter *human pick time* for products in the same aisle, compared to low prevention-focused pickers that would need $15\% * 60 \text{ min} * 6 \text{ hours} = 54 \text{ minutes per day}$ for the same amount of picking in the *human following* setup.
- (3) The average accuracy advantage of *human following* compared to *human leading* setup amounts to an average of $\left(\frac{0.66 \text{ errors}}{20 \text{ min}} \right) * \frac{60 \text{ min}}{20 \text{ min}} * 6 \text{ hours} = 11.88 \text{ fewer errors per order picker per day}$ (Table 7).

6. Conclusion and discussion

This study advances the field of operations management by establishing that *human leading* HRC yields a higher order picking productivity (+8.3% average productivity advantage) and *human following* HRC achieves greater order picking accuracy (−0.66 errors/20 min average accuracy advantage). Furthermore, we identify prevention focus as the key human characteristic that partially explains productivity variations among workers in HRC order picking, with high prevention-focused workers picking up to 40.1% faster than low prevention-focused individuals in the *human following* setup (as long as picks involve travel).

6.1 Theoretical implications

Our findings contribute to the interface between organizational behavior and operations management literature in various ways. Primarily, by empirically comparing human leading vs human following configurations, we demonstrate how perceived autonomy and task interdependence shape both productivity and accuracy. Drawing on SDT (Deci and Ryan, 2012), the human leading setup affords workers high decision autonomy over routing and pacing, thereby satisfying the autonomy need and driving higher productivity. In contrast, the human following setup imposes structured guidance – consistent with competence support in SDT – which mitigates cognitive load at the pick location and thus enhances accuracy. This dual finding extends SDT by showing that autonomy support trades off against competence support in task performance and highlights the importance of designing collaborative systems that balance these psychological needs.

Additionally, by adapting Leader–Member Exchange (LMX) Theory to human–robot dyads, our results suggest that the quality of the leader–follower relationship (whether the human or the robot assumes the leader role directly) affects performance outcomes. When humans lead, autonomous motivation and high-quality exchange with the supportive robot yield greater throughput. When robots lead, clear and consistent signals function as high-quality “LMX” feedback, reducing errors. Thus, we extend LMX Theory by suggesting that non-human agents can fulfill the role of leaders in operational tasks and that the directionality of leadership can fundamentally alter the balance between throughput and accuracy. Finally, although the prevention focus did not change aggregate productivity or accuracy, our detailed analyses reveal that it moderates the productivity gap in humans following these setups. High prevention-focused individuals engage in preparatory behaviors (processing pick information during travel) that offset the structural constraints of following the robot. This finding enriches Regulatory Focus Theory (Higgins, 1998) by demonstrating that task structure interacts with dispositional focus to shape state-dependent performance. It underscores that personality traits exert their strongest effects when behavioral demand (here, interdependence) creates opportunities for vigilance and error-avoidance strategies.

6.2 Practical implications for operations management

Our findings reveal potential trade-offs across different HRC setups, as well as potential strategies to overcome them. Because we use a stylized approach in introducing the leading/following relationship, with the same robot in both setups, our findings are broadly applicable: whether companies are investigating new cobotics solutions, have existing installed setups or use different cobotic setups in different parts of the facility.

As such, for new collaborative robotics solutions, companies can deploy our findings to align setup selection with product characteristics. Specifically, warehouses that need maximum throughput and can justify the higher potential for errors in orders (e.g. spare parts, low-cost products) may benefit from the productivity superiority of a *human leading* setup. Conversely, if higher accuracy is needed (e.g. medicine, high-value products), investing in a *human following* setup may be preferable for reducing the risk of errors, even at the expense of higher throughput. Also, companies that face substantial costs of pick errors due to

increased rework or customer returns (Berger and Ludwig, 2007; Grosse *et al.*, 2013) may especially benefit from a *human following* setup.

Similarly, in existing *human leading* collaborative configurations, companies may reduce the higher potential for errors by introducing quantity reminders or by delaying the sharing of quantity information until the picker reaches the pick location. Furthermore, in *human following* collaborative settings, leveraging the picker's prevention focus can diminish the potential productivity disadvantage. Companies may train workers to develop a higher prevention focus, which can lead to positive effects and better productivity. Alternatively, companies may even evoke picker prevention focus by framing the order picking as a risk-avoidance task to insure against potential productivity loss (Crowe and Higgins, 1997).

6.3 Strengths, limitations and future research directions

Our study's strength lies in two main pillars. First, we conducted a real-effort experiment in a realistic warehouse laboratory environment, using physical robots, a large representative research sample and conditions and reward systems that are reflective of common business practices. This allowed us to draw conclusions with academic rigor, yet directly transferable to practice. We offer a valuable addition to previous literature, which mainly relied on simulated studies, often based on non-empirically validated assumptions. Second, we employed a within-subjects design and objective performance data from two sources (aggregate quality control, detailed robot logs), allowing for both macroscopic and microscopic analysis of performance outcomes and worker behavior. Obviously, a lab experiment also has some limitations.

One limitation is the short duration of our experiment. While we did not find evidence to expect that a longer experimental run would result in different conclusions, a field study would be needed to see if our findings would hold in the longer run. Another limitation of the study is that our sample only included medium to high prevention-focused individuals (ranging from 3 to 5 on a 1–5 scale). As warehouse employees (and even trainees) are trained to have a higher base level of prevention focus to comply with rules and regulations, this might not be just a feature of our sample but a characteristic of the population. One may consider that workers would not even be employed for long if they were truly aloof and did not have a medium level of prevention focus. Another limitation may be the use of low acceleration for the robots due to safety considerations during the experiment, and the use of predefined movements for the robots (which were necessary to ensure comparability between scenarios). In practice, autonomous robots recognize errors and opportunities for optimization and vary their movement, which may subsequently influence the behavior of humans in ways that differ from the ones studied here. Replication field studies in real warehouses may alleviate these concerns. Furthermore, the intertwined nature of framing and spatial positioning in our design means we cannot empirically distinguish which component drives the observed effects. Both elements likely contribute to the outcomes we observed: the verbal framing may influence participants' psychological perception of their role and autonomy, while the spatial positioning creates actual behavioral constraints or freedoms that affect performance. Future research could benefit from factorial designs that systematically vary framing (e.g. "you are leading" vs. "you are following") and spatial positioning (human physically leads vs. follows) independently to isolate their respective contributions to human–robot collaborative performance.

Inevitably, when robots become truly self-sufficient, better and cheaper than humans at most activities, human labor in order picking may become increasingly redundant. Depending on how radically this substitution effect materializes, it has the potential to disrupt and revolutionize the production and operations ecosystem, leading to generalized layoffs with major socioeconomic consequences. This, however, is a limited perspective. The socially responsible alternative for an Industry 5.0 dictates the need for research that reevaluates the threat of job replacement and focuses on solutions that foster harmonious human–machine

coexistence. Therefore, instead of seeking human replacement or crunching humans in robot-paced systems, the path for a sustainable future may lie in the synergistic performance of HRC.

Statement

This research has been approved by the University Ethics Review Board as ETH2122-0634.

References

- 6 River Systems (2019), "6 River Systems explainer", [Video]. YouTube, available at: <https://youtu.be/FFa8zvs2eCw>
- Azadeh, K., De Koster, R. and Roy, D. (2019), "Robotized and automated warehouse systems: review and recent developments", *Transportation Science*, Vol. 53 No. 4, pp. 917-945, (July) doi: [10.1287/trsc.2018.0873](https://doi.org/10.1287/trsc.2018.0873).
- Azadeh, K., Roy, D., de Koster, R. and Khalilabadi, S.M.G. (2025), "Zoning strategies for human-robot collaborative picking", *Decision Sciences*, Vol. 56 No. 1, pp. 50-70, doi: [10.1111/dec.12620](https://doi.org/10.1111/dec.12620).
- Bates, D., Sarkar, D., Bates, M.D. and Matrix, L. (2007), "The lme 4 package", *R Package Version*, Vol. 2 No. 1, p. 74.
- Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014), "Fitting linear mixed-effects models using lme 4", ArXiv Preprint ArXiv: 1406.5823.
- Beersma, B., Homan, A.C., Kleef, G.A.Van and Dreu, C.K.W.De. (2013), "Organizational Behavior and human decision processes outcome interdependence shapes the effects of prevention focus on team processes and performance", *Organizational Behavior and Human Decision Processes*, Vol. 121 No. 2, pp. 194-203, doi: [10.1016/j.obhdp.2013.02.003](https://doi.org/10.1016/j.obhdp.2013.02.003).
- Ben-Ari, M. and Mondada, F. (2017), "Elements of robotics (robots and their applications)", in *Elements of Robotics*, doi: [10.1007/978-3-319-62533-1](https://doi.org/10.1007/978-3-319-62533-1).
- Bendoly, E., Croson, R., Goncalves, P. and Schultz, K. (2010), "Bodies of knowledge for research in behavioral operations", *Production and Operations Management*, Vol. 19 No. 4, pp. 434-452, doi: [10.1111/j.1937-5956.2009.01108.x](https://doi.org/10.1111/j.1937-5956.2009.01108.x).
- Berger, S.M. and Ludwig, T.D. (2007), "Reducing warehouse employee errors using voice-assisted technology that provided immediate feedback", *Journal of Organizational Behavior Management*, Vol. 27 No. 1, pp. 1-31, doi: [10.1300/J075v27n01_01](https://doi.org/10.1300/J075v27n01_01).
- Beuk, F. and Basadur, T. (2016), "Regulatory focus, task engagement and divergent thinking", *Creativity and Innovation Management*, Vol. 25 No. 2, pp. 199-210, doi: [10.1111/caim.12182](https://doi.org/10.1111/caim.12182).
- Bruemmer, D.J., Few, D.A., Boring, R.L., Marble, J.L., Walton, M.C. and Nielsen, C.W. (2005), "Shared understanding for collaborative control", *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 35 No. 4, pp. 494-504, doi: [10.1109/TSMCA.2005.850599](https://doi.org/10.1109/TSMCA.2005.850599).
- Cantor, D.E. and Jin, Y. (2019), "Theoretical and empirical evidence of behavioral and production line factors that influence helping behavior", *Journal of Operations Management*, Vol. 65 No. 4, pp. 312-332, doi: [10.1002/joom.1019](https://doi.org/10.1002/joom.1019).
- Charness, G. and Kuhn, P. (2011), "Chapter 3 - lab labor: what can labor economists learn from the lab?", in Ashenfelter, O. and Card, D. (Eds), *Handbook of Labor Economics, Volume 4, Part A*, Elsevier, pp. 229-330.
- Chernikova, M., Lo Destro, C., Pierro, A., Tory Higgins, E. and Kruglanski, A.W. (2017), "A multilevel analysis of person-group regulatory-mode complementarity: the moderating role of group-task interdependence", *Group Dynamics*, Vol. 21 No. 2, pp. 108-120, doi: [10.1037/gdn0000067](https://doi.org/10.1037/gdn0000067).
- Crowe, E. and Higgins, E.T. (1997), "Regulatory focus and strategic inclinations: promotion and prevention in decision-making", *Organizational Behavior and Human Decision Processes*, Vol. 69 No. 2, pp. 117-132, doi: [10.1006/obhd.1996.2675](https://doi.org/10.1006/obhd.1996.2675).

- De Vries, J., De Koster, R. and Stam, D. (2016a), "Aligning order picking methods, incentive systems, and regulatory focus to increase performance", *Production and Operations Management*, Vol. 25 No. 8, pp. 1363-1376, doi: [10.1111/poms.12547](https://doi.org/10.1111/poms.12547).
- De Vries, J., De Koster, R. and Stam, D. (2016b), "Exploring the role of picker personality in predicting picking performance with pick by voice, pick to light and RF-terminal picking", *International Journal of Production Research*, Vol. 54 No. 8, pp. 2260-2274, doi: [10.1080/00207543.2015.1064184](https://doi.org/10.1080/00207543.2015.1064184).
- De Vries, J., De Koster, R. and Stam, D. (2016c), "Safety does not happen by accident: antecedents to A safer warehouse", *Production and Operations Management*, Vol. 25 No. 8, pp. 1377-1390, doi: [10.1111/poms.12546](https://doi.org/10.1111/poms.12546).
- Deci, E.L. and Ryan, R.M. (2012), "Self-determination theory", in Lange, P.A.M., Kruglanski, A.W. and Higgins, E.T. (Eds), *Handbook of Theories of Social Psychology*, Sage, London, pp. 416-437.
- Deci, E.L., Olafsen, A.H. and Ryan, R.M. (2017), "Self-determination theory in work organizations: the state of a science", *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 4 No. 1, pp. 19-43, doi: [10.1146/annurev-orgpsych-032516-113108](https://doi.org/10.1146/annurev-orgpsych-032516-113108).
- Doerr, K.H., Mitchell, T.R., Schriesheim, C.A., Freed, T. and Zhou, X. (2002), "Heterogeneity and variability in the context of flow lines", *Academy of Management Review*, Vol. 27 No. 4, pp. 594-607, doi: [10.5465/AMR.2002.7566098](https://doi.org/10.5465/AMR.2002.7566098).
- Donohue, K., Özer, Ö. and Zheng, Y. (2020), "Behavioral operations: past, present, and future", *Manufacturing and Service Operations Management*, Vol. 22 No. 1, pp. 191-202, doi: [10.1287/msom.2019.0828](https://doi.org/10.1287/msom.2019.0828).
- Eckerd, S., DuHadway, S., Bendoly, E., Carter, C.R. and Kaufmann, L. (2021), "On making experimental design choices: discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes", *Journal of Operations Management*, Vol. 67 No. 2, pp. 261-275, doi: [10.1002/joom.1128](https://doi.org/10.1002/joom.1128).
- Fragapane, G., De Koster, R., Sgarbossa, F. and Strandhagen, J.O. (2021), "Planning and control of autonomous mobile robots for intralogistics: literature review and research agenda", *European Journal of Operational Research*, Vol. 294 No. 2, pp. 405-426, doi: [10.1016/j.ejor.2021.01.019](https://doi.org/10.1016/j.ejor.2021.01.019).
- Gerstner, C.R. and Day, D.V. (1997), "Meta-analytic review of leader-member exchange theory: correlates and construct issues", *Journal of Applied Psychology*, Vol. 82 No. 6, pp. 827-844, doi: [10.1037//0021-9010.82.6.827](https://doi.org/10.1037//0021-9010.82.6.827).
- Ghorashi Khalilabadi, M., Roy, D. and De Koster, R. (2025), "Picking the best bot: collaboration strategies for humans and bots in order pick systems with traveling salesman problem routing", *Transportation Science*, Vol. 60 No. 1, pp. 101-131, doi: [10.1287/trsc.2024.0969](https://doi.org/10.1287/trsc.2024.0969).
- Globe Newswire (2025), "Netherlands robotics market report 2025-2033", available at: <https://tinyurl.com/2s33ebn8>
- Grosse, E.H., Glock, C.H. and Jaber, M.Y. (2013), "The effect of worker learning and forgetting on storage reassignment decisions in order picking systems", *Computers and Industrial Engineering*, Vol. 66 No. 4, pp. 653-662, doi: [10.1016/j.cie.2013.09.013](https://doi.org/10.1016/j.cie.2013.09.013).
- Grosse, E.H., Glock, C.H. and Neumann, W.P. (2017), "Human factors in order picking: a content analysis of the literature", *International Journal of Production Research*, Vol. 55 No. 5, pp. 1260-1276, doi: [10.1080/00207543.2016.1186296](https://doi.org/10.1080/00207543.2016.1186296).
- Gutelius, B. and Theodore, N. (2019), "The future of warehouse work: technological change in the U.S. Logistics industry", (October), 1-76, available at: <http://laborcenter.berkeley.edu/future-of-warehouse-work/>
- Hagger, M.S., Hardcastle, S.J., Chater, A., Mallett, C., Pal, S. and Chatzisarantis, N.L.D. (2014), "Autonomous and controlled motivational regulations for multiple health-related behaviors: between- and within-participants analyses", *Health Psychology and Behavioral Medicine*, Vol. 2 No. 1, pp. 565-601, doi: [10.1080/21642850.2014.912945](https://doi.org/10.1080/21642850.2014.912945).

- Higgins, E.T. (1998), "Promotion and prevention: regulatory focus as a motivational principle", *Advances in Experimental Social Psychology*, Vol. 30, pp. 1-46, doi: [10.1016/S0065-2601\(08\)60381-0](https://doi.org/10.1016/S0065-2601(08)60381-0).
- Hosseini, Z., Le Blanc, P.M., Demerouti, E., van Gool, P., van den Tooren, M. and Preenen, P.T.Y. (2024), "The impact of working with an automated guided vehicle on boredom and performance: an experimental study in a warehouse environment", *International Journal of Production Research*, pp. 1-21, doi: [10.1080/00207543.2024.2436640](https://doi.org/10.1080/00207543.2024.2436640).
- Inam, R., Raizer, K., Hata, A., Souza, R., Forsman, E., Cao, E. and Wang, S. (2018), "Risk assessment for human-robot collaboration in an automated warehouse scenario", *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA, 2018-Sept(September)*, pp. 743-751, doi: [10.1109/ETFA.2018.8502466](https://doi.org/10.1109/ETFA.2018.8502466).
- International Organization for Standardization (2012), "ISO 8373:2012: robots and robotic devices—vocabulary",
- Lenth, R. (2018), "Emmeans: estimated marginal means, aka least-squares means", *R Package Version*, Vol. 1 No. 1.
- Locus Robotics available at: <https://www.youtube.com/watch?v=3xlm4nSHVx0> (accessed 21 July 2025).
- Löffler, M., Boysen, N. and Schneider, M. (2018), "Picker routing in AGV-assisted order picking systems",
- Lorson, F., Fügener, A. and Hübner, A. (2022), "New team mates in the warehouse: human interactions with automated and robotized systems", *IIEE Transactions*, Vol. 55 No. 5, pp. 536-553, doi: [10.1080/24725854.2022.2072545](https://doi.org/10.1080/24725854.2022.2072545).
- Mangiafico, S. (2020), "Functions to support extension education program evaluation", R Package, R Package, available at: <http://rcompanion.org/depends>
- Matheson, E., Minto, R., Zampieri, E.G.G., Faccio, M. and Rosati, G. (2019), "Human-robot collaboration in manufacturing applications: a review", *Robotics*, Vol. 8 No. 4, pp. 1-25, doi: [10.3390/robotics8040100](https://doi.org/10.3390/robotics8040100).
- Nayyar, A. and Kumar, A. (2020), *A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development*, Springer, Berlin.
- Neumann, W.P., Winkelhaus, S., Grosse, E.H. and Glock, C.H. (2021), "Industry 4.0 and the human factor – a systems framework and analysis methodology for successful development", *International Journal of Production Economics*, Vol. 233, 107992, doi: [10.1016/j.ijpe.2020.107992](https://doi.org/10.1016/j.ijpe.2020.107992).
- Olsen, T.L. and Tomlin, B. (2020), "Industry 4.0: opportunities and challenges for operations management", *Manufacturing and Service Operations Management*, Vol. 22 No. 1, pp. 113-122, doi: [10.1287/msom.2019.0796](https://doi.org/10.1287/msom.2019.0796).
- Pasparakis, A., De Vries, J. and De Koster, R. (2023), "Assessing the impact of human–robot collaborative order picking systems on warehouse workers", *International Journal of Production Research*, Vol. 61 No. 22, pp. 7776-7790, doi: [10.1080/00207543.2023.2183343](https://doi.org/10.1080/00207543.2023.2183343).
- Petković, T., Puljiz, D., Marković, I. and Hein, B. (2019), "Human intention estimation based on hidden Markov model motion validation for safe flexible robotized warehouses", *Robotics and Computer-Integrated Manufacturing*, Vol. 57, pp. 182-196, doi: [10.1016/j.rcim.2018.11.004](https://doi.org/10.1016/j.rcim.2018.11.004).
- Powell, S.G. and Schultz, K.L. (2004), "Throughput in serial lines with state-dependent behavior", *Management Science*, Vol. 50 No. 8, pp. 1095-1105, doi: [10.1287/mnsc.1040.0233](https://doi.org/10.1287/mnsc.1040.0233).
- R Core Team (2020), "R: a language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria. URL, available at: <https://www.R-project.org/>
- Rosenbaum, M.E., Moore, D.L., Cotton, J.L., Cook, M.S., Hieser, R.A., Shovar, M.N. and Gray, M.J. (1980), "Group productivity and process: pure and mixed reward structures and task interdependence", *Journal of Personality and Social Psychology*, Vol. 39 No. 4, pp. 626-642, doi: [10.1037//0022-3514.39.4.626](https://doi.org/10.1037//0022-3514.39.4.626).

- Schriesheim, C.A., Castro, S.L. and Cogliser, C.C. (1999), "Leader-member exchange (LMX) research: a comprehensive review of theory, measurement, and data-analytic practices", *The Leadership Quarterly*, Vol. 10 No. 1, pp. 63-113, doi: [10.1016/s1048-9843\(99\)80009-5](https://doi.org/10.1016/s1048-9843(99)80009-5).
- Schultz, K.L., Juran, D.C., Boudreau, J.W., McClain, J.O. and Thomas, L.J. (1998), "Modeling and worker motivation in JIT production systems", *Management Science*, Vol. 44 No. 12 PART 1, pp. 1595-1607, doi: [10.1287/mnsc.44.12.1595](https://doi.org/10.1287/mnsc.44.12.1595).
- Sgarbossa, F., Grosse, E.H., Neumann, W.P., Battini, D. and Glock, C.H. (2020), "Human factors in production and logistics systems of the future", *Annual Reviews in Control*, Vol. 49, pp. 295-305, doi: [10.1016/j.arcontrol.2020.04.007](https://doi.org/10.1016/j.arcontrol.2020.04.007).
- Skaug, H., Fournier, D., Nielsen, A., Magnusson, A. and Bolker, B. (2014), "GlmADMB: generalized linear mixed models using AD Model Builder", R package version 0.8. 0.
- Statista (2024), "Robotics", available at: <https://www.statista.com/outlook/tmo/robotics/worldwide>
- Steiner, I.D. (1972), *Group Process and Productivity*, Academic Press, New York.
- Tompkins, J.A., White, J.A., Bozer, Y.A. and Tanchoco, J.M.A. (2010), *Facilities Planning*, John Wiley & Sons, Hoboken.
- Toyota-BT available at: <https://www.youtube.com/watch?v=IjdqXAltXgc> (accessed 21 July 2025).
- Tsai, C.Y., Marshall, J.D., Choudhury, A., Serban, A., Hou, Y.T.Y., Jung, M.F., Dionne, S.D. and Yammarino, F.J. (2022), "Human-robot collaboration: a multilevel and integrated leadership framework", *The Leadership Quarterly*, Vol. 33 No. 1, 101594, doi: [10.1016/j.leaqua.2021.101594](https://doi.org/10.1016/j.leaqua.2021.101594).
- Tukey, J.W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, MA, Exploratory Data Analysis. Addison-Wesley, Reading, MA.
- Wallace, C. and Chen, G. (2006), "A multilevel integration of personality, climate, self-regulation, and performance", *Personnel Psychology*, Vol. 59 No. 3, pp. 529-557, doi: [10.1111/j.1744-6570.2006.00046.x](https://doi.org/10.1111/j.1744-6570.2006.00046.x).
- Wang, X.V., Kemény, Z., Váncza, J. and Wang, L. (2017), "Human-robot collaborative assembly in cyber-physical production: classification framework and implementation", *CIRP Annals - Manufacturing Technology*, Vol. 66 No. 1, pp. 5-8, doi: [10.1016/j.cirp.2017.04.101](https://doi.org/10.1016/j.cirp.2017.04.101).
- Wang, L., Liu, S., Liu, H. and Wang, X.V. (2020), "Overview of human-robot collaboration in manufacturing", *Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing*, Springer International Publishing, doi: [10.1007/978-3-030-46212-3](https://doi.org/10.1007/978-3-030-46212-3).
- Wang, Z., Sheu, J., Teo, C. and Xue, G. (2022), "Robot scheduling for mobile-rack warehouses: human-robot coordinated order picking systems", *Production and Operations Management*, Vol. 31 No. 1, pp. 98-116, doi: [10.1111/poms.13406](https://doi.org/10.1111/poms.13406).
- West, B.T., Welch, K.B. and Galecki, A.T. (2014), *Linear Mixed Models: A Practical Guide Using Statistical Software*, Chapman and Hall/CRC, New York.
- Wu, A.C., Gumpertz, M.L., Boos, D.D., Wu, C., Gumpertz, M.L. and Boos, D.D. (2017), "Comparison of GEE, MINQUE, ML, and REML estimating equations for normally distributed data", Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association Stable URL: Comparison of GEE, MI, Vol. 55 No. 2, pp. 125-130, doi: [10.1198/000313001750358608](https://doi.org/10.1198/000313001750358608). available at: <http://www.jstor.org/stable/2685999>
- Yanco, H.A. and Drury, J. (2004), "Classifying human-robot interaction: an updated taxonomy", *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, Vol. 3, pp. 2841-2846.

Corresponding author

René De Koster can be contacted at: rkoster@rsm.nl

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com